Internet Engineering Task Force (IETF)
Request for Comments: 6871
Updates: 5939
Category: Standards Track
ISSN: 2070-1721

R. Gilman
Independent
R. Even
Huawei Technologies
F. Andreasen
Cisco Systems

February 2013

Session Description Protocol (SDP) Media Capabilities Negotiation

Abstract

Session Description Protocol (SDP) capability negotiation provides a general framework for indicating and negotiating capabilities in SDP. The base framework defines only capabilities for negotiating transport protocols and attributes. This documents extends the framework by defining media capabilities that can be used to negotiate media types and their associated parameters.

This document updates the IANA Considerations of RFC 5939.

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 5741.

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/rfc6871.

Table of Contents

1. Introduction

"Session Description Protocol (SDP) Capability Negotiation" [RFC5939] provides a general framework for indicating and negotiating capabilities in SDP [RFC4566]. The base framework defines only capabilities for negotiating transport protocols and attributes.

RFC 5939 [RFC5939] lists some of the issues with the current SDP capability negotiation process. An additional real-life problem is to be able to offer one media stream (e.g., audio) but list the capability to support another media stream (e.g., video) without actually offering it concurrently.

In this document, we extend the framework by defining media capabilities that can be used to indicate and negotiate media types and their associated format parameters. This document also adds the ability to declare support for media streams, the use of which can be offered and negotiated later, and the ability to specify session configurations as combinations of media stream configurations. The definitions of new attributes for media capability negotiation are chosen to make the translation from these attributes to "conventional" SDP [RFC4566] media attributes as straightforward as possible in order to simplify implementation. This goal is intended to reduce processing in two ways: each proposed configuration in an offer may be easily translated into a conventional SDP media stream record for processing by the receiver and the construction of an answer based on a selected proposed configuration would be straightforward.

This document updates RFC 5939 [RFC5939] by updating the IANA considerations. All other extensions defined in this document are considered extensions above and beyond RFC 5939 [RFC5939].

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 RFC 2119 [RFC2119] and indicate requirement levels for compliant implementations.

Actual Configuration: An actual configuration specifies which combinations of SDP session parameters and media stream components can be used in the current offer/answer exchange and with what parameters. Use of an actual configuration does not require any further negotiation in the offer/answer exchange. See RFC 5939 [RFC5939] for further details.

Base Attributes: Conventional SDP attributes appearing in the base configuration of a media block.

Base Configuration: The media configuration represented by a media block exclusive of all the capability negotiation attributes defined in this document, the base capability negotiation document [RFC5939], or any other capability negotiation document. In an offer SDP, the base configuration corresponds to the actual configuration as defined in RFC 5939 [RFC5939].

Conventional Attribute: Any SDP attribute other than those defined by the series of capability negotiation specifications.

Conventional SDP: An SDP record devoid of capability negotiation attributes.

Media Format Capability: A media format, typically a media subtype such as PCMU, H263-1998, or T38, expressed in the form of a capability.

Media Format Parameter Capability: A media format parameter ("a=fmtp" in conventional SDP) expressed in the form of a capability. The media format parameter capability is associated with a media format capability.

Media Capability: The combined set of capabilities associated with expressing a media format and its relevant parameters (e.g., media format parameters and media specific parameters).

Potential Configuration: A potential configuration indicates which combinations of capabilities can be used for the session and its associated media stream components. Potential configurations are not ready for use; however, they are offered for potential use in the current offer/answer exchange. They provide an alternative that may be used instead of the actual configuration, subject to negotiation in the current offer/answer exchange. See RFC 5939 [RFC5939] for further details.

Latent Configuration: A latent configuration indicates which combinations of capabilities could be used in a future negotiation for the session and its associated media stream components. Latent configurations are neither ready for use nor offered for actual or potential use in the current offer/answer exchange. Latent configurations merely inform the other side of possible configurations supported by the entity. Those latent configurations may be used to guide subsequent offer/answer exchanges, but they are not offered for use as part of the current offer/answer exchange.

3. SDP Media Capabilities

The SDP capability negotiation [RFC5939] discusses the use of any SDP [RFC4566] attribute (a=) under the attribute capability "acap". The limitations of using "acap" for "fmtp" and "rtpmap" in a potential configuration are described in RFC 5939 [RFC5939]; for example, they can be used only at the media level since they are media-level attributes. RFC 5939 [RFC5939] does not provide a way to exchange media-level capabilities prior to the actual offer of the associated media stream. This section provides an overview of extensions providing an SDP media capability negotiation solution offering more robust capabilities negotiation. This is followed by definitions of new SDP attributes for the solution and its associated updated offer/answer procedures [RFC3264].

3.1. Requirements

The capability negotiation extensions requirements considered herein are as follows.

REQ-01: Support the specification of alternative (combinations of) media formats (codecs) in a single media block.

REQ-02: Support the specification of alternative media format parameters for each media format.

REQ-03: Retain backward compatibility with conventional SDP. Ensure that each and every offered configuration can be easily translated into a corresponding SDP media block expressed with conventional SDP lines.

REQ-04: Ensure that the scheme operates within the offer/answer model in such a way that media formats and parameters can be agreed upon with a single exchange.

REQ-05: Provide the ability to express offers in such a way that the offerer can receive media as soon as the offer is sent. (Note that the offerer may not be able to render received media prior to exchange of keying material.)

REQ-06: Provide the ability to offer latent media configurations for future negotiation.

REQ-07: Provide reasonable efficiency in the expression of alternative media formats and/or format parameters, especially in those cases in which many combinations of options are offered.

REQ-08: Retain the extensibility of the base capability negotiation mechanism.

REQ-09: Provide the ability to specify acceptable combinations of media streams and media formats. For example, offer a PCMU audio stream with an H264 video stream or a G729 audio stream with an H263 video stream. This ability would give the offerer a means to limit processing requirements for simultaneous streams. This would also permit an offer to include the choice of an audio/T38 stream or an image/T38 stream, but not both.

Other possible extensions have been discussed, but have not been treated in this document. They may be considered in the future. Three such extensions are:

FUT-01: Provide the ability to mix, or change, media types within a single media block. Conventional SDP does not support this capability explicitly; the usual technique is to define a media subtype that represents the actual format within the nominal media type. For example, T.38 FAX as an alternative to audio/PCMU within an audio stream is identified as audio/T38; a separate FAX stream would use image/T38.

FUT-02: Provide the ability to support multiple transport protocols within an active media stream without reconfiguration. This is not explicitly supported by conventional SDP.

FUT-03: Provide capability negotiation attributes for all media-level SDP line types in the same manner as already done for the attribute type, with the exception of the media line type itself. The media line type is handled in a special way to permit compact expression of media coding/format options. The line types are bandwidth ("b="), information ("i="), connection data ("c="), and, possibly, the deprecated encryption key ("k=").

3.2. Solution Overview

The solution consists of new capability attributes corresponding to conventional SDP line types, new parameters for the "pcfg", "acfg", and the new "lcfg" attributes extending the base attributes from RFC 5939 [RFC5939], and a use of the "pcfg" attribute to return capability information in the SDP answer.

Several new attributes are defined in a manner that can be related to the capabilities specified in a media line, and its corresponding "rtpmap" and "fmtp" attributes.

- A new attribute ("a=rmcap") defines RTP-based media format capabilities in the form of a media subtype (e.g., "PCMU"), and its encoding parameters (e.g., "/8000/2"). Each resulting media format type/subtype capability has an associated handle called a media capability number. The encoding parameters are as specified for the "rtpmap" attribute defined in SDP [RFC4566], without the payload type number part.

- A new attribute ("a=omcap") defines other (non-RTP-based) media format capabilities in the form of a media subtype only (e.g., "T38"). Each resulting media format type/subtype capability has an associated handle called a media capability number.

- A new attribute ("a=mfcap") specifies media format parameters associated with one or more media format capabilities. The "mfcap" attribute is used primarily to associate the media format parameters normally carried in the "fmtp" attribute. Note that media format parameters can be used with RTP and non-RTP-based media formats.

- A new attribute ("a=mscap") specifies media parameters associated with one or more media format capabilities. The "mscap" attribute is used to associate capabilities with attributes other than "fmtp" or "rtpmap", for example, the "rtcp-fb" attribute defined in RFC 4585 [RFC4585].

- A new attribute ("a=lcfg") specifies latent media stream configurations when no corresponding media line ("m=") is offered. An example is the offer of latent configurations for video even though no video is currently offered. If the peer indicates support for one or more offered latent configurations, the corresponding media stream(s) may be added via a new offer/answer exchange.

- A new attribute ("a=sescap") is used to specify an acceptable combination of simultaneous media streams and their configurations as a list of potential and/or latent configurations.

New parameters are defined for the potential configuration ("pcfg"), latent configuration ("lcfg"), and accepted configuration ("acfg") attributes to associate the new attributes with particular configurations.

- A new parameter type ("m=") is added to the potential configuration ("a=pcfg:") attribute and the actual configuration ("a=acfg:") attribute defined in RFC 5939 [RFC5939] and to the new latent configuration ("a=lcfg:") attribute. This permits specification of media capabilities (including their associated

parameters) and combinations thereof for the configuration. For example, the "a=pcfg:" line might specify PCMU and telephone events [RFC4733] or G.729B and telephone events as acceptable configurations. The "a=acfg:" line in the answer would specify the configuration chosen.

- o A new parameter type ("pt=") is added to the potential configuration, actual configuration, and latent configuration attributes. This parameter associates RTP payload type numbers with the referenced RTP-based media format capabilities and is appropriate only when the transport protocol uses RTP.

- o A new parameter type ("mt=") is used to specify the media type for latent configurations.

Special processing rules are defined for capability attribute arguments in order to reduce the need to replicate essentially identical attribute lines for the base configuration and potential configurations.

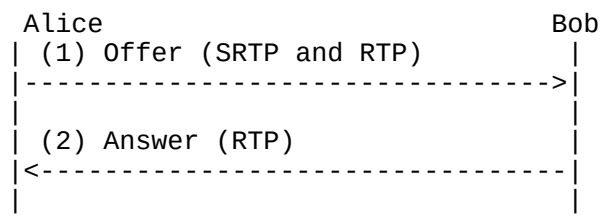
- o A substitution rule is defined for any capability attribute to permit the replacement of the (escaped) media capability number with the media format identifier (e.g., the payload type number in audio/video profiles).

- o Replacement rules are defined for the conventional SDP equivalents of the "mfcap" and "mscap" capability attributes. This reduces the necessity to use the deletion qualifier in the "a=pcfg" parameter in order to ignore "rtpmap", "fmtp", and certain other attributes in the base configuration.

- o An argument concatenation rule is defined for "mfcap" attributes that refer to the same media capability number. This makes it convenient to combine format options concisely by associating multiple mfcap lines with multiple media format capabilities.

This document extends the base protocol extensions to the offer/answer model that allow for capabilities and potential configurations to be included in an offer. Media capabilities constitute capabilities that can be used in potential and latent configurations. Whereas potential configurations constitute alternative offers that may be accepted by the answerer instead of the actual configuration(s) included in the "m=" line(s) and associated parameters, latent configurations merely inform the other side of possible configurations supported by the entity. Those latent configurations may be used to guide subsequent offer/answer exchanges, but they are not part of the current offer/answer exchange.

The mechanism is illustrated by the offer/answer exchange below, where Alice sends an offer to Bob:



Alice's offer includes RTP and Secure Real-time Transport Protocol (SRTP) as alternatives. RTP is the default, but SRTP is the preferred one (long lines are folded to fit the margins):

```
v=0
o=- 25678 753849 IN IP4 192.0.2.1
s=
c=IN IP4 192.0.2.1
t=0 0
a=creq:med-v0
m=audio 3456 RTP/AVP 0 18
a=tcap:1 RTP/SAVP RTP/AVP
a=rtpmap:0 PCMU/8000/1
a=rtpmap:18 G729/8000/1
a=fmtp:18 annexb=yes
a=rmcap:1,4 G729/8000/1
a=rmcap:2 PCMU/8000/1
a=rmcap:5 telephone-event/8000
a=mfcap:1 annexb=no
a=mfcap:4 annexb=yes
a=mfcap:5 0-11
a=acap:1 crypto:1 AES_CM_128_HMAC_SHA1_32 \
inline:NzB4d1BINUAvLEw6UzF3WSJ+PSdFcGdUJShpX1Zj|2^20|1:32
a=pcfg:1 m=4,5|1,5 t=1 a=1 pt=1:100,4:101,5:102
a=pcfg:2 m=2 t=1 a=1 pt=2:103
a=pcfg:3 m=4 t=2 pt=4:18
```

The required base and extensions are provided by the "a=creq" attribute defined in RFC 5939 [RFC5939], with the option tag "med-v0", which indicates that the extension framework defined here must be supported. The base-level capability negotiation support ("cap-v0" [RFC5939]) is implied since it is required for the extensions.

The "m=" line indicates that Alice is offering to use plain RTP with PCMU or G.729B. The media line implicitly defines the default

transport protocol (RTP/AVP in this case) and the default actual configuration.

The "a=tcap:1" line, specified in the SDP capability negotiation base protocol [RFC5939], defines transport protocol capabilities, in this case Secure RTP (SAVP profile) as the first option and RTP (AVP profile) as the second option.

The "a=rmcap:1,4" line defines two G.729 RTP-based media format capabilities, numbered 1 and 4, and their encoding rate. The capabilities are of media type "audio" and subtype G729. Note that the media subtype is explicitly specified here, rather than RTP payload type numbers. This permits the assignment of payload type numbers in the media stream configuration specification. In this example, two G.729 subtype capabilities are defined. This permits the declaration of two sets of formatting parameters for G.729.

The "a=rmcap:2" line defines a G.711 mu-law capability, numbered 2.

The "a=rmcap:5" line defines an audio telephone-event capability, numbered 5.

The "a=mfcap:1" line specifies the "fmtp" formatting parameters for capability 1 (offerer will not accept G.729 Annex B packets).

The "a=mfcap:4" line specifies the "fmtp" formatting parameters for capability 4 (offerer will accept G.729 Annex B packets).

The "a=mfcap:5" line specifies the "fmtp" formatting parameters for capability 5 (the dual-tone multi-frequency (DTMF) touchtones 0-9,*,#).

The "a=acap:1" line specified in the base protocol provides the "crypto" attribute that provides the keying material for SRTP using SDP security descriptions.

The "a=pcfg:" attributes provide the potential configurations included in the offer by reference to the media capabilities, transport capabilities, attribute capabilities, and specified payload type number mappings. Three explicit alternatives are provided; the lowest-numbered one is the preferred one. The "a=pcfg:1 ..." line specifies media capabilities 4 and 5, i.e., G.729B and DTMF (including their associated media format parameters), or media capability 1 and 5, i.e., G.729 and DTMF (including their associated media format parameters). Furthermore, it specifies transport protocol capability 1 (i.e., the RTP/SAVP profile - secure RTP), and the attribute capability 1, i.e., the "crypto" attribute provided. Last, it specifies a payload type number mapping for (RTP-based)

media capabilities 1, 4, and 5, thereby permitting the offerer to distinguish between encrypted media and unencrypted media received prior to receipt of the answer.

Use of unique payload type numbers in alternative configurations is not required; codecs such as Adaptive Multi-Rate Wideband (AMR-WB) [RFC4867] have the potential for so many combinations of options that it may be impractical to define unique payload type numbers for all supported combinations. If unique payload type numbers cannot be specified, then the offerer will be obliged to wait for the SDP answer before rendering received media. For SRTP using Security Descriptions (SDES) inline keying [RFC4568], the offerer will still need to receive the answer before being able to decrypt the stream.

The second alternative ("a=pcfg:2 ...") specifies media capability 2, i.e., PCMU, under the RTP/SAVP profile, with the same SRTP key material.

The third alternative ("a=pcfg:3 ...") offers G.729B unsecured; its only purpose in this example is to show a preference for G.729B over PCMU.

Per RFC 5939 [RFC5939], the media line, with any qualifying attributes such as "fmtp" or "rtpmap", is itself considered a valid configuration (the current actual configuration); it has the lowest preference (per RFC 5939 [RFC5939]).

Bob receives the SDP offer from Alice. Bob supports G.729B, PCMU, and telephone events over RTP, but not SRTP, hence he accepts the potential configuration 3 for RTP provided by Alice. Bob generates the following answer:

```
v=0
o=- 24351 621814 IN IP4 192.0.2.2
s=
c=IN IP4 192.0.2.2
t=0 0
a=csup:med-v0
m=audio 4567 RTP/AVP 18
a=rtpmap:18 G729/8000
a=fmtp:18 annexb=yes
a=acfg:3 m=4 t=2 pt=4:18
```

Bob includes the "a=csup" and "a=acfg" attributes in the answer to inform Alice that he can support the med-v0 level of capability negotiations. Note that in this particular example, the answerer supported the capability extensions defined here; however, had he not, he would simply have processed the offer based on the offered

PCMU and G.729 codecs under the RTP/AVP profile only. Consequently, the answer would have omitted the "a=csup" attribute line and chosen one or both of the PCMU and G.729 codecs instead. The answer carries the accepted configuration in the "m=" line along with corresponding "rtpmap" and/or "fmtp" parameters, as appropriate.

Note that per the base protocol, after the above, Alice MAY generate a new offer with an actual configuration ("m=" line, etc.) corresponding to the actual configuration referenced in Bob's answer (not shown here).

3.3. New Capability Attributes

In this section, we present the new attributes associated with indicating the media capabilities for use by the SDP capability negotiation. The approach taken is to keep things similar to the existing media capabilities defined by the existing media descriptions ("m=" lines) and the associated "rtpmap" and "fmtp" attributes. We use media subtypes and "media capability numbers" to link the relevant media capability parameters. This permits the capabilities to be defined at the session level and be used for multiple streams, if desired. For RTP-based media formats, payload types are then specified at the media level (see Section 3.3.4.2).

A media capability merely indicates possible support for the media type and media format(s) and parameters in question. In order to actually use a media capability in an offer/answer exchange, it MUST be referenced in a potential configuration.

Media capabilities, i.e., the attributes associated with expressing media capability formats, parameters, etc., can be provided at the session level and/or the media level. Media capabilities provided at the session level may be referenced in any "pcfg" or "lcfg" attribute at the media level (consistent with the media type), whereas media capabilities provided at the media level may be referenced only by the "pcfg" or "lcfg" attribute within that media stream. In either case, the scope of the <med-cap-num> is the entire session description. This enables each media capability to be uniquely referenced across the entire session description (e.g., in a potential configuration).

3.3.1. The Media Format Capability Attributes

Media subtypes can be expressed as media format capabilities by use of the "a=rmcap" and "a=omcap" attributes. The "a=rmcap" attribute MUST be used for RTP-based media, whereas the "a=omcap" attribute MUST be used for non-RTP-based (other) media formats. The two attributes are defined as follows:

```
a=rmcap:<media-cap-num-list> <encoding-name>/<clock-rate>
                             [/<encoding-parms>]

a=omcap:<media-cap-num-list> <format-name>
```

where <media-cap-num-list> is a (list of) media capability number(s) used to number a media format capability, the <encoding name> or <format-name> is the media subtype, e.g., H263-1998, PCMU, or T38, <clock rate> is the encoding rate, and <encoding parms> are the media encoding parameters for the media subtype. All media format capabilities in the list are assigned to the same media type/subtype. Each occurrence of the "rmcap" and "omcap" attribute MUST use unique values in their <media-cap-num-list>; the media capability numbers are shared between the two attributes and the numbers MUST be unique across the entire SDP session. In short, the "rmcap" and "omcap" attributes define media format capabilities and associate them with a media capability number in the same manner as the "rtpmap" attribute defines them and associates them with a payload type number. Additionally, the attributes allow multiple capability numbers to be defined for the media format in question by specifying a range of media capability numbers. This permits the media format to be associated with different media parameters in different configurations. When a range of capability numbers is specified, the first (leftmost) capability number MUST be strictly smaller than the second (rightmost), i.e., the range increases and covers at least two numbers.

In ABNF [RFC5234], we have:

```
media-capability-line = rtp-mcap / non-rtp-mcap

rtp-mcap           = "a=rmcap:" media-cap-num-list
                        1*WSP encoding-name "/" clock-rate
                        ["/" encoding-parms]
non-rtp-mcap       = "a=omcap:" media-cap-num-list 1*WSP format-name
media-cap-num-list = media-cap-num-element
                     *("," media-cap-num-element)
media-cap-num-element = media-cap-num
                              / media-cap-num-range
media-cap-num-range = media-cap-num "-" media-cap-num
media-cap-num      = NonZeroDigit *9(DIGIT)
encoding-name      = token ;defined in RFC 4566
clock-rate         = NonZeroDigit *9(DIGIT)
encoding-parms     = token
format-name        = token ;defined in RFC 4566
NonZeroDigit       = %x31-39    ; 1-9
```

The encoding-name, clock-rate, and encoding-params are as defined to appear in an "rtpmap" attribute for each media type/subtype. Thus, it is easy to convert an "rmcap" attribute line into one or more "rtpmap" attribute lines, once a payload type number is assigned to a media-cap-num (see Section 3.3.5).

The format-name is a media format description for non-RTP-based media as defined for the <fmt> part of the media description ("m=" line) in SDP [RFC4566]. In simple terms, it is the name of the media format, e.g., "t38". This form can also be used in cases such as Binary Floor Control Protocol (BFCP) [RFC4585] where the fmt list in the "m=" line is effectively ignored (BFCP uses "*").

The "rmcap" and "omcap" attributes can be provided at the session level and/or the media level. There can be more than one "rmcap" and more than one "omcap" attribute at both the session and media levels (i.e., more than one of each at the session level and more than one of each in each media description). Media capability numbers cannot include leading zeroes, and each media-cap-num MUST be unique within the entire SDP record; it is used to identify that media capability in potential, latent, and actual configurations, and in other attribute lines as explained below. Note that the media-cap-num values are shared between the "rmcap" and "omcap" attributes; hence, the uniqueness requirement applies to the union of them. When the media capabilities are used in a potential, latent, or actual configuration, the media formats referred by those configurations apply at the media level, irrespective of whether the media capabilities themselves were specified at the session or media level. In other words, the media capability applies to the specific media description associated with the configuration that invokes it.

For example:

```
v=0
o=- 24351 621814 IN IP4 192.0.2.2
s=
c=IN IP4 192.0.2.2
t=0 0
a=rmcap:1 L16/8000/1
a=rmcap:2 L16/16000/2
a=rmcap:3 H263-1998/90000
a=omcap:4 example
m=audio 54320 RTP/AVP 0
a=pcfg:1 m=1|2, pt=1:99,2:98
m=video 66544 RTP/AVP 100
a=rtpmap:100 H264/90000
a=pcfg:10 m=3 pt=3:101
a=tcap:1 TCP
a=pcfg:11 m=4 t=1
```

3.3.2. The Media Format Parameter Capability Attribute

This attribute is used to associate media format specific parameters with one or more media format capabilities. The form of the attribute is

```
a=mfcap:<media-caps> <list of parameters>
```

where <media-caps> permits the list of parameters to be associated with one or more media format capabilities and the format parameters are specific to the type of media format. The mfcap lines map to a single traditional SDP "fmtp" attribute line (one for each entry in <media-caps>) of the form

```
a=fmtp:<fmt> <list of parameters>
```

where <fmt> is the media format parameter defined in RFC 4566 [RFC4566], as appropriate for the particular media stream. The "mfcap" attribute MUST be used to encode attributes for media capabilities, which would conventionally appear in an "fmtp" attribute. The existing "acap" attribute MUST NOT be used to encode "fmtp" attributes.

The "mfcap" attribute adheres to SDP [RFC4566] attribute production rules with

```
media-format-parameter-capability =
       "a=mfcap:" media-cap-num-list 1*WSP fmt-specific-param-list
fmt-specific-param-list = text ; defined in RFC 4566
```

Note that media format parameters can be used with RTP-based and non-RTP-based media formats.

3.3.2.1. Media Format Parameter Concatenation Rule

The appearance of media subtypes with a large number of formatting options (e.g., AMR-WB [RFC4867]), coupled with the restriction that only a single "fmtp" attribute can appear per media format, suggests that it is useful to create a combining rule for "mfcap" parameters that are associated with the same media capability number. Therefore, different mfcap lines MAY include the same media-cap-num in their media-cap-num-list. When a particular media capability is selected for processing, the parameters from each mfcap line that references the particular capability number in its media-cap-num-list are concatenated together via ";", in the order the "mfcap" attributes appear in the SDP record, to form the equivalent of a single "fmtp" attribute line. This permits one to define a separate mfcap line for a single parameter and value that is to be applied to each media capability designated in the media-cap-num-list. This provides a compact method to specify multiple combinations of format parameters when using codecs with multiple format options. Note that order-dependent parameters SHOULD be placed in a single mfcap line to avoid possible problems with line rearrangement by a middlebox.

Format parameters are not parsed by SDP; their content is specific to the media type/subtype. When format parameters for a specific media capability are combined from multiple "a=mfcap" lines that reference that media capability, the format-specific parameters are concatenated together and separated by ";" for construction of the corresponding format attribute ("a=fmtp"). The resulting format attribute will look something like the following (without line breaks):

```
a=fmtp:<fmt> <fmt-specific-param-list1>;
             <fmt-specific-param-list2>;
             ...
```

where <fmt> depends on the transport protocol in the manner defined in RFC 4566 [RFC4566]. SDP cannot assess the legality of the resulting parameter list in the "a=fmtp" line; the user must take care to ensure that legal parameter lists are generated.

The "mfcap" attribute can be provided at the session level and the media level. There can be more than one "mfcap" attribute at the session or media level. The unique media-cap-num is used to associate the parameters with a media capability.

As a simple example, a G.729 capability is, by default, considered to support comfort noise as defined by Annex B. Capabilities for G.729 with and without comfort noise support may thus be defined by:

```
a=rmcap:1,2 G729/8000
a=mfcap:2 annexb:no
```

Media capability 1 supports G.729 with Annex B, whereas media capability 2 supports G.729 without Annex B.

Example for H.263 video:

```
a=rmcap:1 H263-1998/90000
a=rmcap:2 H263-2000/90000
a=mfcap:1 CIF=4;QCIF=2;F=1;K=1
a=mfcap:2 profile=2;level=2.2
```

Finally, for six format combinations of the Adaptive Multi-Rate codec:

```
a=rmcap:1-3 AMR/8000/1
a=rmcap:4-6 AMR-WB/16000/1
a=mfcap:1,2,3,4 mode-change-capability=1
a=mfcap:5,6 mode-change-capability=2
a=mfcap:1,2,3,5 max-red=220
a=mfcap:3,4,5,6 octet-align=1
a=mfcap:1,3,5 mode-set=0,2,4,7
a=mfcap:2,4,6 mode-set=0,3,5,6
```

So that AMR codec #1, when specified in a "pcfg" attribute within an audio stream block (and assigned payload type number 98) as in:

```
a=pcfg:1 m=1 pt=1:98
```

is essentially equivalent to the following:

```
m=audio 49170 RTP/AVP 98
a=rtpmap:98 AMR/8000/1
a=fmtp:98 mode-change-capability=1; \
max-red=220; mode-set=0,2,4,7
```

and AMR codec #4 with payload type number 99, depicted by the potential configuration:

```
a=pcfg:4 m=4, pt=4:99
```

is equivalent to the following:

```
m=audio 49170 RTP/AVP 99
a=rtpmap:99 AMR-WB/16000/1
a=fmtp:99 mode-change-capability=1; octet-align=1; \
mode-set=0,3,5,6
```

and so on for the other four combinations. SDP could thus convert the media capabilities specifications into one or more alternative media stream specifications, one of which can be chosen for the answer.

3.3.3. The Media-Specific Capability Attribute

Attributes and parameters associated with a media format are typically specified using the "rtpmap" and "fmtp" attributes in SDP, and the similar "rmcap" and "mfcap" attributes in SDP media capabilities. Some SDP extensions define other attributes that need to be associated with media formats, for example, the "rtcp-fb" attribute defined in RFC 4585 [RFC4585]. Such media-specific attributes, beyond the "rtpmap" and "fmtp" attributes, may be associated with media capability numbers via a new media-specific attribute, "mscap", of the following form:

```
a=mscap:<media caps star> <att field> <att value>
```

where <media caps star> is a (list of) media capability number(s), <att field> is the attribute name, and <att value> is the value field for the named attribute. Note that the media capability numbers refer to media format capabilities specified elsewhere in the SDP ("rmcap" and/or "omcap"). If a range of capability numbers is specified, the first (leftmost) capability number MUST be strictly smaller than the second (rightmost). The media capability numbers may include a wildcard ("*"), which will be used instead of any payload type mappings in the resulting SDP (see, e.g., RFC 4585 [RFC4585] and the example below). In ABNF, we have:

```
media-specific-capability = "a=mscap:"
                             media-caps-star
                             1*WSP att-field ; from RFC 4566
                             1*WSP att-value ; from RFC 4566
media-caps-star           =  media-cap-star-element
                               *("," media-cap-star-element)
media-cap-star-element    = (media-cap-num [wildcard])
                          / (media-cap-num-range [wildcard])
wildcard                  = "*"
```

Given an association between a media capability and a payload type number as specified by the "pt=" parameters in a "pcfg" attribute line, a mscap line may be translated easily into a conventional SDP attribute line of the form:

```
a=<att field>":"<fmt> <att value> ; <fmt> defined in SDP [RFC4566]
```

A resulting attribute that is not a legal SDP attribute, as specified by RFC 4566, MUST be ignored by the receiver.

If a media capability number (or range) contains a wildcard character at the end, any payload type mapping specified for that media-specific capability (or range of capabilities) will use the wildcard character in the resulting SDP instead of the payload type specified in the payload type mapping ("pt" parameter) in the configuration attribute.

A single mscap line may refer to multiple media capabilities by use of a capability number range; this is equivalent to multiple mscap lines, each with the same attribute values (but different media capability numbers), one line per media capability.

Multiple mscap lines may refer to the same media capability, but, unlike the "mfcap" attribute, no concatenation operation is defined. Hence, multiple mscap lines applied to the same media capability are equivalent to multiple lines of the specified attribute in a conventional media record.

Here is an example with the "rtcp-fb" attribute, modified from an example in RFC 5104 [RFC5104] (with the session level and audio media omitted). If the offer contains a media block like the following (note the wildcard character),

```
m=video 51372 RTP/AVP 98
a=rtpmap:98 H263-1998/90000
a=tcap:1 RTP/AVPF
a=rmcap:1 H263-1998/90000
a=mscap:1 rtcp-fb ccm tstr
a=mscap:1 rtcp-fb ccm fir
a=mscap:1* rtcp-fb ccm tmmbr smaxpr=120
a=pcfg:1 t=1 m=1 pt=1:98
```

and if the proposed configuration is chosen, then the equivalent media block would look like the following

```
m=video 51372 RTP/AVPF 98
a=rtpmap:98 H263-1998/90000
a=rtcp-fb:98 ccm tstr
a=rtcp-fb:98 ccm fir
a=rtcp-fb:* ccm tmmbr smaxpr=120
```

3.3.4. New Configuration Parameters

Along with the new attributes for media capabilities, new extension parameters are defined for use in the potential configuration, the actual configuration, and/or the new latent configuration defined in Section 3.3.5.

3.3.4.1. The Media Configuration Parameter (m=)

The media configuration parameter is used to specify the media format(s) and related parameters for a potential, actual, or latent configuration. Adhering to the ABNF for extension-config-list in RFC 5939 [RFC5939] with

```
ext-cap-name = "m"
ext-cap-list = media-cap-num-list
               [*(BAR media-cap-num-list)]
```

we have

```
media-config-list = ["+"] "m=" media-cap-num-list
                                *(BAR media-cap-num-list)
                    ;BAR is defined in RFC 5939
                    ;media-cap-num-list is defined above
```

Alternative media configurations are separated by a vertical bar ("|"). The alternatives are ordered by preference, most-preferred first. When media capabilities are not included in a potential configuration at the media level, the media type and media format from the associated "m=" line will be used. The use of the plus sign ("+") is described in RFC 5939.

3.3.4.2. The Payload Type Number Mapping Parameter (pt=)

The payload type number mapping parameter is used to specify the payload type number to be associated with each RTP-based media format in a potential, actual, or latent configuration. We define the payload type number mapping parameter, payload-number-config-list, in accordance with the extension-config-list format defined in RFC 5939 [RFC5939]. In ABNF:

```
payload-number-config-list = ["+"] "pt=" media-map-list
media-map-list      = media-map *("," media-map)
media-map           = media-cap-num ":" payload-type-number
                        ; media-cap-num is defined in Section 3.3.1
payload-type-number = NonZeroDigit *2(DIGIT) ; RTP payload
                                             ; type number
```

The example in Section 3.3.7 shows how the parameters from the rmcap line are mapped to payload type numbers from the "pcfg" "pt" parameter. The use of the plus sign ("+") is described in RFC 5939 [RFC5939].

A latent configuration represents a future capability; hence, the "pt=" parameter is not directly meaningful in the "lcfg" attribute because no actual media session is being offered or accepted. It is permitted in order to tie any payload type number parameters within attributes to the proper media format. A primary example is the case of format parameters for the Redundant Audio Data (RED) [RFC2198] payload, which are payload type numbers. Specific payload type numbers used in a latent configuration MAY be interpreted as suggestions to be used in any future offer based on the latent configuration, but they are not binding; the offerer and/or answerer may use any payload type numbers each deems appropriate. The use of explicit payload type numbers for latent configurations can be avoided by use of the parameter substitution rule of Section 3.3.7. Future extensions are also permitted. Note that leading zeroes are not permitted.

3.3.4.3. The Media Type Parameter

When a latent configuration is specified (always at the media level), indicating the ability to support an additional media stream, it is necessary to specify the media type (audio, video, etc.) as well as the format and transport type. The media type parameter is defined in ABNF as

```
media-type = ["+"] "mt=" media; media defined in RFC 4566
```

At present, the media-type parameter is used only in the latent configuration attribute, and the use of the "+" prefix to specify that the entire attribute line is to be ignored if the mt= parameter is not understood is unnecessary. However, if the media-type parameter is later added to an existing capability attribute such as "pcfg", then the "+" would be useful. The media format(s) and transport type(s) are specified using the media configuration parameter ("+m=") defined above, and the transport parameter ("t=") defined in RFC 5939 [RFC5939], respectively.

3.3.5. The Latent Configuration Attribute

One of the goals of this work is to permit the exchange of supportable media configurations in addition to those offered or accepted for immediate use. Such configurations are referred to as "latent configurations". For example, a party may offer to establish a session with an audio stream, and, at the same time, announce its ability to support a video stream as part of the same session. The offerer can supply its video capabilities by offering one or more latent video configurations along with the media stream for audio; the responding party may indicate its ability and willingness to support such a video session by returning a corresponding latent configuration.

Latent configurations returned in SDP answers MUST match offered latent configurations (or parameter subsets thereof). Therefore, it is appropriate for the offering party to announce most, if not all, of its capabilities in the initial offer. This choice has been made in order to keep the size of the answer more compact by not requiring acap, rmcap, tcap, etc. lines in the answer.

Latent configurations may be announced by use of the latent configuration attribute, which is defined in a manner very similar to the potential configuration attribute. The latent configuration attribute combines the properties of a media line and a potential configuration. A latent configuration MUST include a media type (mt=) and a transport protocol configuration parameter since the latent configuration is independent of any media line present. In most cases, the media configuration (m=) parameter needs to be present as well (see Section 4 for examples). The "lcfg" attribute is a media-level attribute.

> The "lcfg" attribute is defined as a media-level attribute since it specifies a possible future media stream. However, the "lcfg" attribute is not necessarily related to the media description within which it is provided. Session capability attributes ("a=sescap") may be used to indicate supported media stream configurations.

Each media line in an SDP description represents an offered simultaneous media stream, whereas each latent configuration represents an additional stream that may be negotiated in a future offer/answer exchange. Session capability attributes may be used to determine whether a latent configuration may be used to form an offer for an additional simultaneous stream or to reconfigure an existing stream in a subsequent offer/answer exchange.

The latent configuration attribute is of the form:

```
a=lcfg:<config-number> <latent-cfg-list>
```

which adheres to the SDP [RFC4566] "attribute" production with att-field and att-value defined as:

```
att-field  = "lcfg"
att-value  = config-number 1*WSP lcfg-cfg-list
config-number = NonZeroDigit *9(DIGIT)  ;DIGIT defined in RFC 5234
lcfg-cfg-list = media-type 1*WSP pot-cfg-list
                              ; as defined in RFC 5939
                              ; and extended herein
```

The media-type (mt=) parameter identifies the media type (audio, video, etc.) to be associated with the latent media stream, and it MUST be present. The pot-cfg-list MUST contain a transport-protocol-config-list (t=) parameter and a media-config-list (m=) parameter. The pot-cfg-list MUST NOT contain more than one instance of each type of parameter list. As specified in RFC 5939 [RFC5939], the use of the "+" prefix with a parameter indicates that the entire configuration MUST be ignored if the parameter is not understood; otherwise, the parameter itself may be ignored.

Media stream payload numbers are not assigned by a latent configuration. Assignment will take place if and when the corresponding stream is actually offered via an "m=" line in a later exchange. The payload-number-config-list is included as a parameter to the "lcfg" attribute in case it is necessary to tie payload numbers in attribute capabilities to specific media capabilities.

If an "lcfg" attribute invokes an "acap" attribute that appears at the session level, then that attribute will be expected to appear at the session level of a subsequent offer when and if a corresponding media stream is offered. Otherwise, "acap" attributes that appear at the media level represent media-level attributes. Note, however, that "rmcap", omcap, "mfcap", "mscap", and "tcap" attributes may appear at the session level because they always result in media-level attributes or "m=" line parameters.

The configuration numbers for latent configurations do not imply a preference; the offerer will imply a preference when actually offering potential configurations derived from latent configurations negotiated earlier. Note, however, that the offerer of latent configurations MAY specify preferences for combinations of potential and latent configurations by use of the "sescap" attribute defined in Section 3.3.8. For example, if an SDP offer contains, say, an audio

stream with "pcfg:1", and two latent video configurations, "lcfg:2" and "lcfg:3", then a session with one audio stream and one video stream could be specified by including "a=sescap:1 1,2|3". One audio stream and two video streams could be specified by including "a=sescap:2 1,2,3" in the offer. In order to permit combinations of latent and potential configurations in session capabilities, latent configuration numbers MUST be different from those used for potential configurations. This restriction is especially important if the offerer does not require cmed-v0 capability and the recipient of the offer doesn't support it. If the "lcfg" attribute is not recognized, the capability attributes intended to be associated with it may be confused with those associated with a potential configuration of some other media stream. Note also that leading zeroes are not permitted in configuration numbers.

If a cryptographic attribute, such as the SDES "a=crypto:" attribute [RFC4568], is referenced by a latent configuration through an "acap" attribute, any keying material required in the conventional attribute, such as the SDES key/salt string, MUST be included in order to satisfy formatting rules for the attribute. Since the keying material will be visible but not actually used at this stage (since it's a latent configuration), the value(s) of the keying material MUST NOT be a real value used for real exchange of media, and the receiver of the "lcfg" attribute MUST ignore the value(s).

3.3.6. Enhanced Potential Configuration Attribute

The present work requires new extensions (parameters) for the "pcfg" attribute defined in the SDP capability negotiation base protocol [RFC5939]. The parameters and their definitions are "borrowed" from the definitions provided for the latent configuration attribute in Section 3.3.5. The expanded ABNF definition of the "pcfg" attribute is

```
a=pcfg: <config-number> [<pot-cfg-list>]
```

where

```
config-number = 1*DIGIT ;defined in [RFC5234]
pot-cfg-list  = pot-config *(1*WSP pot-config)
pot-config    =  attribute-config-list / ;def in [RFC5939]
     transport-protocol-config-list / ;defined in [RFC5939]
     extension-config-list / ;[RFC5939]
     media-config-list / ; Section 3.3.4.1
     payload-number-config-list ; Section 3.3.4.2
```

Except for the extension-config-list, the pot-cfg-list MUST NOT contain more than one instance of each parameter list.

3.3.6.1. Returning Capabilities in the Answer

Potential and/or latent configuration attributes may be returned within an answer SDP to indicate the ability of the answerer to support alternative configurations of the corresponding stream(s). For example, an offer may include multiple potential configurations for a media stream and/or latent configurations for additional streams. The corresponding answer will indicate (via an "acfg" attribute) the configuration accepted and used to construct the base configuration for each active media stream in the reply, but the reply MAY also contain potential and/or latent configuration attributes, with parameters, to indicate which other offered configurations would be acceptable. This information is useful if it becomes desirable to reconfigure a media stream, e.g., to reduce resource consumption.

When potential and/or latent configurations are returned in an answer, all numbering MUST refer to the configuration and capability attribute numbering of the offer. The offered capability attributes need not be returned in the answer. The answer MAY include additional capability attributes and/or configurations (with distinct numbering). The parameter values of any returned "pcfg" or "lcfg" attributes MUST be a subset of those included in the offered configurations and/or those added by the answerer; values MAY be omitted only if they were indicated as alternative sets, or optional, in the original offer. The parameter set indicated in the returned "acfg" attribute need not be repeated in a returned "pcfg" attribute. The answerer MAY return more than one "pcfg" attribute with the same configuration number if it is necessary to describe selected combinations of optional or alternative parameters.

Similarly, one or more session capability attributes ("a=sescap") MAY be returned to indicate which of the offered session capabilities is/are supportable by the answerer (see Section 3.3.8).

Note that, although the answerer MAY return capabilities beyond those included by the offerer, these capabilities MUST NOT be used to form any base level media description in the answer. For this reason, it is advisable for the offerer to include most, if not all, potential and latent configurations it can support in the initial offer, unless the size of the resulting SDP is a concern. Either party MAY later announce additional capabilities by renegotiating the session in a second offer/answer exchange.

3.3.6.2. Payload Type Number Mapping

When media format capabilities defined in "rmcap" attributes are used in potential configuration lines, the transport protocol uses RTP and it is necessary to assign payload type numbers. In some cases, it is desirable to assign different payload type numbers to the same media format capability when used in different potential configurations. One example is when configurations for AVP and SAVP are offered: the offerer would like the answerer to use different payload type numbers for encrypted and unencrypted media, so the offerer can decide whether or not to render early media that arrives before the answer is received.

> For example, if use of AVP was selected by the answerer, then media received by the offerer is not encrypted; hence, it can be played out prior to receiving the answer. Conversely, if SAVP was selected, cryptographic parameters and keying material present in the answer may be needed to decrypt received media. If the offer configuration indicated that AVP media uses one set of payload types and SAVP a different set, then the offerer will know whether media received prior to the answer is encrypted or not by simply looking at the RTP payload type number in the received packet.

This association of distinct payload type number(s) with different transport protocols requires a separate pcfg line for each protocol. Clearly, this technique cannot be used if the number of potential configurations exceeds the number of possible payload type numbers.

3.3.6.3. Processing of Media-Format-Related Conventional Attributes for Potential Configurations

When media capabilities negotiation is employed, SDP records are likely to contain conventional attributes such as "rtpmap", "fmtp", and other media-format-related lines, as well as capability attributes such as "rmcap", omcap, "mfcap", and "mscap" that map into those conventional attributes when invoked by a potential configuration. In such cases, it MAY be appropriate to employ the delete-attributes option [RFC5939] in the attribute configuration list parameter in order to avoid the generation of conflicting "fmtp" attributes for a particular configuration. Any media-specific attributes in the media block that refer to media formats not used by the potential configuration MUST be ignored.

For example:

```
v=0
o=- 25678 753849 IN IP4 192.0.2.1
s=
c=IN IP4 192.0.2.1
t=0 0
a=creq:med-v0
m=audio 3456 RTP/AVP 0 18 100
a=rtpmap:100 telephone-event
a=fmtp:100 0-11
a=rmcap:1 PCMU/8000
a=rmcap:2 G729/8000
a=rmcap:3 telephone-event/8000
a=mfcap:3 0-15
a=pcfg:1 m=2,3|1,3 a=-m pt=1:0,2:18,3:100
a=pcfg:2
```

In this example, PCMU is media capability 1, G729 is media capability 2, and telephone-event is media capability 3. The a=pcfg:1 line specifies that the preferred configuration is G.729 with extended DTMF events, second is G.711 mu-law with extended DTMF events, and the base media-level attributes are to be deleted. Intermixing of G.729, G.711, and "commercial" DTMF events is least preferred (the base configuration provided by the "m=" line, which is, by default, the least preferred configuration). The "rtpmap" and "fmtp" attributes of the base configuration are replaced by the "rmcap" and "mfcap" attributes when invoked by the proposed configuration.

If the preferred configuration is selected, the SDP answer will look like the following

```
v=0
o=- 25678 753849 IN IP4 192.0.2.1
s=
c=IN IP4 192.0.2.1
t=0 0
a=csup:med-v0
m=audio 3456 RTP/AVP 18 100
a=rtpmap:100 telephone-event/8000
a=fmtp:100 0-15
a=acfg:1 m=2,3 pt=1:0,2:18,3:100
```

3.3.7. Substitution of Media Payload Type Numbers in Capability Attribute Parameters

In some cases, for example, when an RFC 2198 [RFC2198] redundancy audio subtype (RED) capability is defined in an "mfcap" attribute, the parameters to an attribute may contain payload type numbers. Two options are available for specifying such payload type numbers. They may be expressed explicitly, in which case they are bound to actual payload types by means of the payload type number parameter (pt=) in the appropriate potential or latent configuration. For example, the following SDP fragment defines a potential configuration with redundant G.711 mu-law

```
m=audio 45678 RTP/AVP 0
a=rtpmap:0 PCMU/8000
a=rmcap:1 PCMU/8000
a=rmcap:2 RED/8000
a=mfcap:2 0/0
a=pcfg:1 m=2,1 pt=2:98,1:0
```

The potential configuration is then equivalent to

```
m=audio 45678 RTP/AVP 98 0
a=rtpmap:0 PCMU/8000
a=rtpmap:98 RED/8000
a=fmtp:98 0/0
```

A more general mechanism is provided via the parameter substitution rule. When an "mfcap", "mscap", or "acap" attribute is processed, its arguments will be scanned for a payload type number escape sequence of the following form (in ABNF):

```
ptn-esc = "%m=" media-cap-num "%" ; defined in Section 3.3.1
```

If the sequence is found, the sequence is replaced by the payload type number assigned to the media capability number, as specified by the "pt=" parameter in the selected potential configuration; only actual payload type numbers are supported -- wildcards are excluded. The sequence "%%" (null digit string) is replaced by a single percent sign and processing continues with the next character, if any.

For example, the above offer sequence could have been written as

```
m=audio 45678 RTP/AVP 0
a=rtpmap:0 PCMU/8000
a=rmcap:1 PCMU/8000
a=rmcap:2 RED/8000
a=mfcap:2 %m=1%/%m=1%
a=pcfg:1 m=2,1 pt=2:98,1:0
```

and the equivalent SDP is the same as above.

3.3.8. The Session Capability Attribute

Potential and latent configurations enable offerers and answerers to express a wide range of alternative configurations for current and future negotiation. However, in practice, it may not be possible to support all combinations of these configurations.

The session capability attribute provides a means for the offerer and/or the answerer to specify combinations of specific media stream configurations that it is willing and able to support. Each session capability in an offer or answer MAY be expressed as a list of required potential configurations, and MAY include a list of optional potential and/or latent configurations.

The choices of session capabilities may be based on processing load, total bandwidth, or any other criteria of importance to the communicating parties. If the answerer supports media capabilities negotiation, and session configurations are offered, it MUST accept one of the offered configurations, or it MUST refuse the session. Therefore, if the offer includes any session capabilities, it SHOULD include all the session capabilities the offerer is willing to support.

The session capability attribute is a session-level attribute described by

"a=sescap:" <session num> <list of configs>

which corresponds to the standard value attribute definition with

```
att-field        = "sescap"
att-value        = session-num 1*WSP list-of-configs
                   [1*WSP optional-configs]
session-num      = NonZeroDigit *9(DIGIT)  ; DIGIT defined
                                           ; in RFC 5234
list-of-configs  = alt-config *("," alt-config)
optional-configs = "[" list-of-configs "]"
alt-config       = config-number *("|" config-number)
```

The session-num identifies the session: a lower-number session is preferred over a higher-number session, and leading zeroes are not permitted. Each alt-config list specifies alternative media configurations within the session; preference is based on config-num as specified in RFC 5939 [RFC5939]. Note that the session preference order, when present, takes precedence over the individual media stream configuration preference order.

Use of session capability attributes requires that configuration numbers assigned to potential and latent configurations MUST be unique across the entire session; RFC 5939 [RFC5939] requires only that "pcfg" configuration numbers be unique within a media description. Also, leading zeroes are not permitted.

As an example, consider an endpoint that is capable of supporting an audio stream with either one H.264 video stream or two H.263 video streams with a floor control stream. In the latter case, the second video stream is optional. The SDP offer might look like the following (offering audio, an H.263 video streams, BFCP and another optional H.263 video stream) -- the empty lines are added for readability only (not part of valid SDP):

```
v=0
o=- 25678 753849 IN IP4 192.0.2.1
s=
c=IN IP4 192.0.2.1
t=0 0
a=creq:med-v0
a=sescap:2 1,2,5,[3]
a=sescap:1 1,4

m=audio 54322 RTP/AVP 0
a=rtpmap:0 PCMU/8000
a=pcfg:1
```

```
m=video 22344 RTP/AVP 102
a=rtpmap:102 H263-1998/90000
a=fmtp:102 CIF=4;QCIF=2;F=1;K=1
i=main video stream
a=label:11
a=pcfg:2
a=rmcap:1 H264/90000
a=mfcap:1 profile-level-id=42A01E; packetization-mode=2
a=acap:1 label:13
a=pcfg:4 m=1 a=1 pt=1:104

m=video 33444 RTP/AVP 103
a=rtpmap:103 H263-1998/90000
a=fmtp:103 CIF=4;QCIF=2;F=1;K=1
i=secondary video (slides)
a=label:12
a=pcfg:3

m=application 33002 TCP/BFCP *
a=setup:passive
a=connection:new
a=floorid:1 m-stream:11 12
a=floor-control:s-only
a=confid:4321
a=userid:1234
a=pcfg:5
```

If the answerer understands MediaCapNeg, but cannot support the Binary Floor Control Protocol, then it would respond with (invalid empty lines in SDP included again for readability):

```
v=0
o=- 25678 753849 IN IP4 192.0.2.1
s=
c=IN IP4 192.0.2.22
t=0 0
a=csup:med-v0
a=sescap:1 1,4

m=audio 23456 RTP/AVP 0
a=rtpmap:0 PCMU/8000
a=acfg:1

m=video 41234 RTP/AVP 104
a=rtpmap:104 H264/90000
a=fmtp:104 profile-level-id=42A01E; packetization-mode=2
a=acfg:4 m=1 a=1 pt=1:104
```

```
m=video 0 RTP/AVP 103
a=acfg:3

m=application 0 TCP/BFCP *
a=acfg:5
```

An endpoint that doesn't support media capabilities negotiation, but does support H.263 video, would respond with one or two H.263 video streams. In the latter case, the answerer may issue a second offer to reconfigure the session to one audio and one video channel using H.264 or H.263.

Session capabilities can include latent capabilities as well. Here's a similar example in which the offerer wishes to initially establish an audio stream, and prefers to later establish two video streams with chair control. If the answerer doesn't understand Media CapNeg, or cannot support the dual video streams or flow control, then it may support a single H.264 video stream. Note that establishment of the most favored configuration will require two offer/answer exchanges.

```
v=0
o=- 25678 753849 IN IP4 192.0.2.1
s=
c=IN IP4 192.0.2.1
t=0 0
a=creq:med-v0
a=sescap:1 1,3,4,5
a=sescap:2 1,2
a=sescap:3 1

a=rmcap:1 H263-1998/90000
a=mfcap:1 CIF=4;QCIF=2;F=1;K=1
a=tcap:1 RTP/AVP TCP/BFCP
m=audio 54322 RTP/AVP 0
a=rtpmap:0 PCMU/8000
a=pcfg:1
m=video 22344 RTP/AVP 102
a=rtpmap:102 H264/90000
a=fmtp:102 profile-level-id=42A01E; packetization-mode=2
a=label:11
a=content:main
a=pcfg:2
a=lcfg:3 mt=video t=1 m=1 a=31,32
a=acap:31 label:12
a=acap:32 content:main
a=lcfg:4 mt=video t=1 m=1 a=41,42
a=acap:41 label:13
a=acap:42 content:slides
```

```
a=lcfg:5 mt=application m=51 t=51
a=tcap:51 TCP/BFCP
a=omcap:51 *
a=acap:51 setup:passive
a=acap:52 connection:new
a=acap:53 floorid:1 m-stream:12 13
a=acap:54 floor-control:s-only
a=acap:55 confid:4321
a=acap:56 userid:1234
```

In this example, the default offer, as seen by endpoints that do not understand capabilities negotiation, proposes a PCMU audio stream and an H.264 video stream. Note that the offered lcfg lines for the video streams don't carry "pt=" parameters because they're not needed (payload type numbers will be assigned in the offer/answer exchange that establishes the streams). Note also that the three "rmcap", "mfcap", and "tcap" attributes used by "lcfg:3" and "lcfg:4" are included at the session level so they may be referenced by both latent configurations. As per Section 3.3, the media attributes generated from the "rmcap", "mfcap", and "tcap" attributes are always media-level attributes. If the answerer supports Media CapNeg, and supports the most desired configuration, it would return the following SDP:

```
v=0
o=- 25678 753849 IN IP4 192.0.2.1
s=
c=IN IP4 192.0.2.22
t=0 0
a=csup:med-v0
a=sescap:1 1,3,4,5
a=sescap:2 1,2
a=sescap:3 1
m=audio 23456 RTP/AVP 0
a=rtpmap:0 PCMU/8000
a=acfg:1
m=video 0 RTP/AVP 102
a=pcfg:2
a=lcfg:3 mt=video t=1 m=1 a=31,32
a=lcfg:4 mt=video t=1 m=1 a=41,42
a=lcfg:5 mt=application t=2
```

This exchange supports immediate establishment of an audio stream for preliminary conversation. This exchange would presumably be followed at the appropriate time with a "reconfiguration" offer/answer exchange to add the video and chair control streams.

3.4. Offer/Answer Model Extensions

In this section, we define extensions to the offer/answer model defined in RFC 3264 [RFC3264] and RFC 5939 [RFC5939] to allow for media format and associated parameter capabilities, latent configurations, and acceptable combinations of media stream configurations to be used with the SDP capability negotiation framework. Note that the procedures defined in this section extend the offer/answer procedures defined in RFC 5939 [RFC5939] Section 6; those procedures form a baseline set of capability negotiation offer/answer procedures that MUST be followed, subject to the extensions defined here.

SDP capability negotiation [RFC5939] provides a relatively compact means to offer the equivalent of an ordered list of alternative configurations for offered media streams (as would be described by separate "m=" lines and associated attributes). The attributes "acap", "mscap", "mfcap", "omcap", and "rmcap" are designed to map somewhat straightforwardly into equivalent "m=" lines and conventional attributes when invoked by a "pcfg", "lcfg", or "acfg" attribute with appropriate parameters. The "a=pcfg:" lines, along with the "m=" line itself, represent offered media configurations. The "a=lcfg:" lines represent alternative capabilities for future use.

3.4.1. Generating the Initial Offer

The media capabilities negotiation extensions defined in this document cover the following categories of features:

- Media format capabilities and associated parameters ("rmcap", "omcap", "mfcap", and "mscap" attributes)
- Potential configurations using those media format capabilities and associated parameters
- Latent media streams ("lcfg" attribute)
- Acceptable combinations of media stream configurations ("sescap" attribute).

The high-level description of the operation is as follows:

When an endpoint generates an initial offer and wants to use the functionality described in the current document, it SHOULD identify and define the media formats and associated parameters it can support via the "rmcap", "omcap", "mfcap", and "mscap" attributes. The SDP media line(s) ("m=") should be made up with the actual configuration

to be used if the other party does not understand capability negotiations (by default, this is the least preferred configuration). Typically, the media line configuration will contain the minimum acceptable configuration from the offerer's point of view.

Preferred configurations for each media stream are identified following the media line. The present offer may also include latent configuration ("lcfg") attributes, at the media level, describing media streams and/or configurations the offerer is not now offering but that it is willing to support in a future offer/answer exchange. A simple example might be the inclusion of a latent video configuration in an offer for an audio stream.

Lastly, if the offerer wishes to impose restrictions on the combinations of potential configurations to be used, it will include session capability ("sescap") attributes indicating those.

If the offerer requires the answerer to understand the media capability extensions, the offerer MUST include a "creq" attribute containing the value "med-v0". If media capability negotiation is required only for specific media descriptions, the "med-v0" value MUST be provided only in "creq" attributes within those media descriptions, as described in RFC 5939 [RFC5939].

Below, we provide a more detailed description of how to construct the offer SDP.

3.4.1.1. Offer with Media Capabilities

For each RTP-based media format the offerer wants to include as a media format capability, the offer MUST include an "rmcap" attribute for the media format as defined in Section 3.3.1.

For each non-RTP-based media format the offer wants to include as a media format capability, the offer MUST include an "omcap" attribute for the media format as defined in Section 3.3.1.

Since the media capability number space is shared between the "rmcap" and "omcap" attributes, each media capability number provided (including ranges) MUST be unique in the entire SDP.

If an "fmtp" parameter value is needed for a media format (whether or not it is RTP based) in a media capability, then the offer MUST include one or more "mfcap" parameters with the relevant "fmtp" parameter values for that media format as defined in Section 3.3.2. When multiple "mfcap" parameters are provided for a given media capability, they MUST be provided in accordance with the concatenation rules in Section 3.3.2.1.

For each of the media format capabilities above, the offer MAY include one or more "mscap" parameters with attributes needed for those specific media formats as defined in Section 3.3.3. Such attributes will be instantiated at the media level; hence, session-level-only attributes MUST NOT be used in the "mscap" parameter. The "mscap" parameter MUST NOT include an "rtpmap" or "fmtp" attribute ("rmcap" and "mfcap" are used instead).

If the offerer wants to limit the relevance (and use) of a media format capability or parameter to a particular media stream, the media format capability or parameter MUST be provided within the corresponding media description. Otherwise, the media format capabilities and parameters MUST be provided at the session level. Note, however, that the attribute or parameter embedded in these will always be instantiated at the media level.

> This is due to those parameters being effectively media-level parameters. If session-level attributes are needed, the "acap" attribute defined in RFC 5939 [RFC5939] can be used; however, it does not provide for media-format-specific instantiation.

Inclusion of the above does not constitute an offer to use the capabilities; a potential configuration is needed for that. If the offerer wants to offer one or more of the media capabilities above, they MUST be included as part of a potential configuration ("pcfg") attribute as defined in Section 3.3.4. Each potential configuration MUST include a config-number, and each config-number MUST be unique in the entire SDP (note that this differs from RFC 5939 [RFC5939], which only requires uniqueness within a media description). Also, the config-number MUST NOT overlap with any config-number used by a latent configuration in the SDP. As described in RFC 5939 [RFC5939], lower config-numbers indicate a higher preference; the ordering still applies within a given media description only though.

For a media capability to be included in a potential configuration, there MUST be an "m=" parameter in the "pcfg" attribute referencing the media capability number in question. When one or more media capabilities are included in an offered potential configuration ("pcfg"), they completely replace the list of media formats offered in the actual configuration ("m=" line). Any attributes included for those formats remain in the SDP though (e.g., "rtpmap", "fmtp", etc.). For non-RTP-based media formats, the format-name (from the "omcap" media capability) is simply added to the "m=" line as a media format (e.g., t38). For RTP-based media, payload type mappings MUST be provided by use of the "pt" parameter in the potential configuration (see Section 3.3.4.2); payload type escaping may be used in "mfcap", "mscap", and "acap" attributes as defined in Section 3.3.7.

Note that the "mt" parameter MUST NOT be used with the "pcfg" attribute (since it is defined for the "lcfg" attribute only); the media type in a potential configuration cannot be changed from that of the encompassing media description.

3.4.1.2. Offer with Latent Configuration

If the offerer wishes to offer one or more latent configurations for future use, the offer MUST include a latent configuration attribute ("lcfg") for each as defined in Section 3.3.6.

Each "lcfg" attribute

- MUST be specified at the media level
- MUST include a config-number that is unique in the entire SDP (including for any potential configuration attributes). Note that config-numbers in latent configurations do not indicate any preference order
- MUST include a media type ("mt")
- MUST reference a valid transport capability ("t")

Each "lcfg" attribute MAY include additional capability references, which may refer to capabilities anywhere in the session description, subject to any restrictions normally associated with such capabilities. For example, a media-level attribute capability must be present at the media level in some media description in the SDP. Note that this differs from the potential configuration attribute, which cannot validly refer to media-level capabilities in another media description (per RFC 5939 [RFC5939], Section 3.5.1).

> Potential configurations constitute an actual offer and may instantiate a referenced capability. Latent configurations are not actual offers; hence, they cannot instantiate a referenced capability. Therefore, it is safe for those to refer to capabilities in another media description.

3.4.1.3. Offer with Configuration Combination Restrictions

If the offerer wants to indicate restrictions or preferences among combinations of potential and/or latent configurations, a session capability ("sescap") attribute MUST be provided at the session level for each such combination as described in Section 3.3.8. Each "sescap" attribute MUST include a session-num that is unique in the

entire SDP; the lower the session-num the more preferred that combination is. Furthermore, "sescap" preference order takes precedence over any order specified in individual "pcfg" attributes.

For example, if we have pcfg-1 and pcfg-2, and sescap-1 references pcfg-2, whereas sescap-2 references pcfg-1, then pcfg-2 will be the most preferred potential configuration. Without the sescap, pcfg-1 would be the most preferred.

3.4.2. Generating the Answer

When receiving an offer, the answerer MUST check the offer for "creq" attributes containing the value "med-v0"; answerers compliant with this specification will support this value in accordance with the procedures specified in RFC 5939 [RFC5939].

The SDP MAY contain

- Media format capabilities and associated parameters ("rmcap", "omcap", "mfcap", and "mscap" attributes)
- Potential configurations using those media format capabilities and associated parameters
- Latent media streams ("lcfg" attribute)
- Acceptable combinations of media stream configurations ("sescap" attribute)

The high-level informative description of the operation is as follows:

When the answering party receives the offer, if it supports the required capability negotiation extensions, it should select the most-preferred configuration it can support for each media stream, and build its answer accordingly. The configuration selected for each accepted media stream is placed into the answer as a media line with associated parameters and attributes. If a proposed configuration is chosen for a given media stream, the answer must contain an actual configuration ("acfg") attribute for that media stream to indicate which offered "pcfg" attribute was used to build the answer. The answer should also include any potential or latent configurations the answerer can support, especially any configurations compatible with other potential or latent configurations received in the offer. The answerer should make note of those configurations it might wish to offer in the future.

Below we provide a more detailed normative description of how the answerer processes the offer SDP and generates an answer SDP.

3.4.2.1. Processing Media Capabilities and Potential Configurations

The answerer MUST first determine if it needs to perform media capability negotiation by examining the SDP for valid and preferred potential configuration attributes that include media configuration parameters (i.e., an "m" parameter in the "pcfg" attribute).

Such a potential configuration is valid if

1. It is valid according to the rules defined in RFC 5939 [RFC5939].

2. It contains a config-number that is unique in the entire SDP and does not overlap with any latent configuration config-numbers.

3. All media format capabilities ("rmcap" or "omcap"), media format parameter capabilities ("mfcap"), and media-specific capabilities ("mscap") referenced by the potential configuration ("m" parameter) are valid themselves (as defined in Sections 3.3.1, 3.3.2, and 3.3.3) and each of them is provided either at the session level or within this particular media description.

4. All RTP-based media format capabilities ("rmcap") have a corresponding payload type ("pt") parameter in the potential configuration that results in mapping to a valid payload type that is unique within the resulting SDP.

5. Any concatenation (see Section 3.3.2.1) and substitution (see Section 3.3.7) applied to any capability ("mfcap", "mscap", or "acap") referenced by this potential configuration results in a valid SDP.

Note that since SDP does not interpret the value of "fmtp" parameters, any resulting "fmtp" parameter value will be considered valid.

Secondly, the answerer MUST determine the order in which potential configurations are to be negotiated. In the absence of any session capability ("sescap") attributes, this simply follows the rules of RFC 5939 [RFC5939], with a lower config-number within a media description being preferred over a higher one. If a valid "sescap" attribute is present, the preference order provided in the "sescap" attribute MUST take precedence. A "sescap" attribute is considered valid if

1. It adheres to the rules provided in Section 3.3.8.

2. All the configurations referenced by the "sescap" attribute are valid themselves (note that this can include the actual, potential, and latent configurations).

The answerer MUST now process the offer for each media stream based on the most preferred valid potential configuration in accordance with the procedures specified in RFC 5939 [RFC5939], Section 3.6.2, and further extended below:

- If one or more media format capabilities are included in the potential configuration, then they replace all media formats provided in the "m=" line for that media description. For non-RTP-based media formats ("omcap"), the format-name is added. For RTP-based media formats ("rmcap"), the payload-type specified in the payload-type mapping ("pt") is added and a corresponding "rtpmap" attribute is added to the media description.

- If one or more media format parameter capabilities are included in the potential configuration, then the corresponding "fmtp" attributes are added to the media description. Note that this inclusion is done indirectly via the media format capability.

- If one or more media-specific capabilities are included in the potential configuration, then the corresponding attributes are added to the media description. Note that this inclusion is done indirectly via the media format capability.

- When checking to see if the answerer supports a given potential configuration that includes one or more media format capabilities, the answerer MUST support at least one of the media formats offered. If he does not, the answerer MUST proceed to the next potential configuration based on the preference order that applies.

- If session capability ("sescap") preference ordering is included, then the potential configuration selection process MUST adhere to the ordering provided. Note that this may involve coordinated selection of potential configurations between media descriptions. The answerer MUST accept one of the offered sescap combinations (i.e., all the required potential configurations specified) or it MUST reject the entire session.

Once the answerer has selected a valid and supported offered potential configuration for all of the media streams (or has fallen back to the actual configuration plus any added session attributes), the answerer MUST generate a valid answer SDP as described in RFC 5939 [RFC5939], Section 3.6.2, and further extended below:

- Additional answer capabilities and potential configurations MAY be returned in accordance with Section 3.3.6.1. Capability numbers and configuration numbers for those MUST be distinct from the ones used in the offer SDP.

- Latent configuration processing and answer generation MUST be performed, as specified below.

- Session capability specification for the potential and latent configurations in the answer MAY be included (see Section 3.3.8).

3.4.2.2. Latent Configuration Processing

The answerer MUST determine if it needs to perform any latent configuration processing by examining the SDP for valid latent configuration attributes ("lcfg"). An "lcfg" attribute is considered valid if:

- It adheres to the description in Section 3.3.5.

- It includes a config-number that is unique in the entire SDP and does not overlap with any potential configuration config-number.

- It includes a valid media type ("mt=").

- It references a valid transport capability ("t=").

- All other capabilities referenced by it are valid.

For each such valid latent configuration in the offer, the answerer checks to see if it could support the latent configuration in a subsequent offer/answer exchange. If so, it includes the latent configuration with the same configuration number in the answer, similar to the way potential configurations are processed and the selected one returned in an actual configuration attribute (see RFC 5939 [RFC5939]). If the answerer supports only a (non-mandatory) subset of the parameters offered in a latent configuration, the answer latent configuration will include only those parameters supported (similar to "acfg" processing). Note that latent configurations do not constitute an actual offer at this point in time; they merely indicate additional configurations that could be supported.

If a session capability ("sescap") attribute is included and it references a latent configuration, then the answerer processing of that latent configuration must be done within the constraints specified by that session capability. That is, it must be possible to support it at the same time as any required (i.e., non-optional) potential configurations in the session capability. The answerer may in turn add his own sescap indications in the answer as well.

3.4.3. Offerer Processing of the Answer

The offerer MUST process the answer in accordance with Section 3.6.3 of RFC 5939 [RFC5939] and the further explanation below.

When the offerer processes the answer SDP based on a valid actual configuration attribute in the answer, and that valid configuration includes one or more media capabilities, the processing MUST furthermore be done as if the offer was sent using those media capabilities instead of the actual configuration. In particular, the media formats in the "m=" line, and any associated payload type mappings ("rtpmap"), "fmtp" parameters ("mfcap"), and media-specific attributes ("mscap") MUST be used. Note that this may involve use of concatenation and substitution rules (see Sections 3.3.2.1 and 3.3.7). The actual configuration attribute may also be used to infer the lack of acceptability of higher-preference configurations that were not chosen, subject to any constraints provided by a session capability ("sescap") attribute in the offer. Note that the SDP capability negotiation base specification [RFC5939] requires the answerer to choose the highest-preference configuration it can support, subject to local policies.

When the offerer receives the answer, it SHOULD furthermore make note of any capabilities and/or latent configurations included for future use, and any constraints on how those may be combined.

3.4.4. Modifying the Session

If, at a later time, one of the parties wishes to modify the operating parameters of a session, e.g., by adding a new media stream, or by changing the properties used on an existing stream, it can do so via the mechanisms defined for offer/answer [RFC3264]. If the initiating party has remembered the codecs, potential configurations, latent configurations, and session capabilities provided by the other party in the earlier negotiation, it MAY use this knowledge to maximize the likelihood of a successful modification of the session. Alternatively, the initiator MAY perform a new capabilities exchange as part of the reconfiguration.

In such a case, the new capabilities will replace the previously negotiated capabilities. This may be useful if conditions change on the endpoint.

4. Examples

In this section, we provide examples showing how to use the media capabilities with the SDP capability negotiation.

4.1. Alternative Codecs

This example provides a choice of one of six variations of the Adaptive Multi-Rate codec. In this example, the default configuration as specified by the media line is the same as the most preferred configuration. Each configuration uses a different payload type number so the offerer can interpret early media.

```
v=0
o=- 25678 753849 IN IP4 192.0.2.1
s=
c=IN IP4 192.0.2.1
t=0 0
a=creq:med-v0
m=audio 54322 RTP/AVP 96
a=rtpmap:96 AMR-WB/16000/1
a=fmtp:96 mode-change-capability=1; max-red=220; \
mode-set=0,2,4,7
a=rmcap:1,3,5 audio AMR-WB/16000/1
a=rmcap:2,4,6 audio AMR/8000/1
a=mfcap:1,2,3,4 mode-change-capability=1
a=mfcap:5,6 mode-change-capability=2
a=mfcap:1,2,3,5 max-red=220
a=mfcap:3,4,5,6 octet-align=1
a=mfcap:1,3,5 mode-set=0,2,4,7
a=mfcap:2,4,6 mode-set=0,3,5,6
a=pcfg:1 m=1 pt=1:96
a=pcfg:2 m=2 pt=2:97
a=pcfg:3 m=3 pt=3:98
a=pcfg:4 m=4 pt=4:99
a=pcfg:5 m=5 pt=5:100
a=pcfg:6 m=6 pt=6:101
```

In the above example, media capability 1 could have been excluded from the first "rmcap" declaration and from the corresponding "mfcap" attributes, and the "pcfg:1" attribute line could have been simply "pcfg:1".

The next example offers a video stream with three options of H.264 and four transports. It also includes an audio stream with different audio qualities: four variations of AMR, or AC3. The offer looks something like the following:

```
v=0
o=- 25678 753849 IN IP4 192.0.2.1
s=An SDP Media NEG example
c=IN IP4 192.0.2.1
t=0 0
a=creq:med-v0
a=ice-pwd:speEc3QGZiNWpVLFJhQX
m=video 49170 RTP/AVP 100
c=IN IP4 192.0.2.56
a=maxprate:1000
a=rtcp:51540
a=sendonly
a=candidate 12345 1 UDP 9 192.0.2.56 49170 host
a=candidate 23456 2 UDP 9 192.0.2.56 51540 host
a=candidate 34567 1 UDP 7 198.51.100.1 41345 srflx raddr \
192.0.2.56 rport 49170
a=candidate 45678 2 UDP 7 198.51.100.1 52567 srflx raddr \
192.0.2.56 rport 51540
a=candidate 56789 1 UDP 3 192.0.2.100 49000 relay raddr \
192.0.2.56 rport 49170
a=candidate 67890 2 UDP 3 192.0.2.100 49001 relay raddr \
192.0.2.56 rport 51540
b=AS:10000
b=TIAS:10000000
b=RR:4000
b=RS:3000
a=rtpmap:100 H264/90000
a=fmtp:100 profile-level-id=42A01E; packetization-mode=2; \
sprop-parameter-sets=Z0IACpZTBYmI,aMljiA==; \
sprop-interleaving-depth=45; sprop-deint-buf-req=64000; \
sprop-init-buf-time=102478; deint-buf-cap=128000
a=tcap:1 RTP/SAVPF RTP/SAVP RTP/AVPF
a=rmcap:1-3,7-9 H264/90000
a=rmcap:4-6 rtx/90000
a=mfcap:1-9 profile-level-id=42A01E
a=mfcap:1-9 aMljiA==
a=mfcap:1,4,7 packetization-mode=0
a=mfcap:2,5,8 packetization-mode=1
a=mfcap:3,6,9 packetization-mode=2
a=mfcap:1-9 sprop-parameter-sets=Z0IACpZTBYmI
a=mfcap:1,7 sprop-interleaving-depth=45; \
sprop-deint-buf-req=64000; sprop-init-buf-time=102478; \
deint-buf-cap=128000
```

```
a=mfcap:4 apt=100
a=mfcap:5 apt=99
a=mfcap:6 apt=98
a=mfcap:4-6 rtx-time=3000
a=mscap:1-6 rtcp-fb nack
a=acap:1 crypto:1 AES_CM_128_HMAC_SHA1_80 \
inline:d0RmdmcmVCspeEc3QGZiNWpVLFJhQX1cfHAwJSoj|220|1:32
a=pcfg:1 t=1 m=1,4 a=1 pt=1:100,4:97
a=pcfg:2 t=1 m=2,5 a=1 pt=2:99,4:96
a=pcfg:3 t=1 m=3,6 a=1 pt=3:98,6:95
a=pcfg:4 t=2 m=7 a=1 pt=7:100
a=pcfg:5 t=2 m=8 a=1 pt=8:99
a=pcfg:6 t=2 m=9 a=1 pt=9:98
a=pcfg:7 t=3 m=1,3 pt=1:100,4:97
a=pcfg:8 t=3 m=2,4 pt=2:99,4:96
a=pcfg:9 t=3 m=3,6 pt=3:98,6:95
m=audio 49176 RTP/AVP 101 100 99 98
c=IN IP4 192.0.2.56
a=ptime:60
a=maxptime:200
a=rtcp:51534
a=sendonly
a=candidate 12345 1 UDP 9 192.0.2.56 49176 host
a=candidate 23456 2 UDP 9 192.0.2.56 51534 host
a=candidate 34567 1 UDP 7 198.51.100.1 41348 srflx \
raddr 192.0.2.56 rport 49176
a=candidate 45678 2 UDP 7 198.51.100.1 52569 srflx \
raddr 192.0.2.56 rport 51534
a=candidate 56789 1 UDP 3 192.0.2.100 49002 relay \
raddr 192.0.2.56 rport 49176
a=candidate 67890 2 UDP 3 192.0.2.100 49003 relay \
raddr 192.0.2.56 rport 51534
b=AS:512
b=TIAS:512000
b=RR:4000
b=RS:3000
a=maxprate:120
a=rtpmap:98 AMR-WB/16000
a=fmtp:98 octet-align=1; mode-change-capability=2
a=rtpmap:99 AMR-WB/16000
a=fmtp:99 octet-align=1; crc=1; mode-change-capability=2
a=rtpmap:100 AMR-WB/16000/2
a=fmtp:100 octet-align=1; interleaving=30
a=rtpmap:101 AMR-WB+/72000/2
a=fmtp:101 interleaving=50; int-delay=160000;
a=rmcap:14 ac3/48000/6
a=acap:23 crypto:1 AES_CM_128_HMAC_SHA1_80 \
inline:d0RmdmcmVCspeEc3QGZiNWpVLFJhQX1cfHAwJSoj|220|1:32
```

```
a=tcap:4 RTP/SAVP
a=pcfg:10 t=4 a=23
a=pcfg:11 t=4 m=14 a=23 pt=14:102
```

This offer illustrates the advantage in compactness that arises if one can avoid deleting the base configuration attributes and recreating them in "acap" attributes for the potential configurations.

4.2. Alternative Combinations of Codecs (Session Configurations)

If an endpoint has limited signal processing capacity, it might be capable of supporting, say, a G.711 mu-law audio stream in combination with an H.264 video stream, or a G.729B audio stream in combination with an H.263-1998 video stream. It might then issue an offer like the following:

```
v=0
o=- 25678 753849 IN IP4 192.0.2.1
s=
c=IN IP4 192.0.2.1
t=0 0
a=creq:med-v0
a=sescap:1 2,4
a=sescap:2 1,3
m=audio 54322 RTP/AVP 18
a=rtpmap:18 G729/8000
a=fmtp:18 annexb=yes
a=rmcap:1 PCMU/8000
a=pcfg:1 m=1 pt=1:0
a=pcfg:2
m=video 54344 RTP/AVP 100
a=rtpmap:100 H263-1998/90000
a=rmcap:2 H264/90000
a=mfcap:2 profile-level-id=42A01E; packetization-mode=2
a=pcfg:3 m=2 pt=2:101
a=pcfg:4
```

Note that the preferred session configuration (and the default as well) is G.729B with H.263. This overrides the individual media stream preferences that are PCMU and H.264 by the potential configuration numbering rule.

4.3. Latent Media Streams

Consider a case in which the offerer can support either G.711 mu-law or G.729B, along with DTMF telephony events for the 12 common touchtone signals, but is willing to support simple G.711 mu-law

audio as a last resort. In addition, the offerer wishes to announce its ability to support video and Message Session Relay Protocol (MSRP) in the future, but does not wish to offer a video stream or an MSRP stream at present. The offer might look like the following:

```
v=0
o=- 25678 753849 IN IP4 192.0.2.1
s=
c=IN IP4 192.0.2.1
t=0 0
a=creq:med-v0
m=audio 23456 RTP/AVP 0
a=rtpmap:0 PCMU/8000
a=rmcap:1 PCMU/8000
a=rmcap:2 G729/8000
a=rmcap:3 telephone-event/8000
a=mfcap:3 0-11
a=pcfg:1 m=1,3|2,3 pt=1:0,2:18,3:100
a=lcfg:2 mt=video t=1 m=10|11
a=rmcap:10 H263-1998/90000
a=rmcap:11 H264/90000
a=tcap:1 RTP/AVP
a=lcfg:3 mt=message t=2 m=20
a=tcap:2 TCP/MSRP
a=omcap:20 *
```

The first "lcfg" attribute line ("lcfg:2") announces support for H.263 and H.264 video (H.263 preferred) for future negotiation. The second "lcfg" attribute line ("lcfg:3") announces support for MSRP for future negotiation. The "m=" line and the "rtpmap" attribute offer an audio stream and provide the lowest precedence configuration (PCMU without any DTMF encoding). The rmcap lines define the RTP-based media format capabilities (PCMU, G729, telephone-event, H263-1998, and H264) and the omcap line defines the non-RTP-based media format capability (wildcard). The "mfcap" attribute provides the format parameters for telephone-event, specifying the 12 commercial DTMF 'digits'. The "pcfg" attribute line defines the most-preferred media configuration as PCMU plus DTMF events and the next-most-preferred configuration as G.729B plus DTMF events.

If the answerer is able to support all the potential configurations, and also support H.263 video (but not H.264), it would reply with an answer like the following:

```
v=0
o=- 24351 621814 IN IP4 192.0.2.2
s=
c=IN IP4 192.0.2.2
t=0 0
a=csup:med-v0
m=audio 54322 RTP/AVP 0 100
a=rtpmap:0 PCMU/8000
a=rtpmap:100 telephone-event/8000
a=fmtp:100 0-11
a=acfg:1 m=1,3 pt=1:0,3:100
a=pcfg:1 m=2,3 pt=2:18,3:100
a=lcfg:2 mt=video t=1 m=10
```

The "lcfg" attribute line announces the capability to support H.263 video at a later time. The media line and subsequent "rtpmap" and "fmtp" attribute lines present the selected configuration for the media stream. The "acfg" attribute line identifies the potential configuration from which it was taken, and the "pcfg" attribute line announces the potential capability to support G.729 with DTMF events as well. If, at some later time, congestion becomes a problem in the network, either party may, with expectation of success, offer a reconfiguration of the media stream to use G.729 in order to reduce packet sizes.

5. IANA Considerations

5.1. New SDP Attributes

IANA has registered the following new SDP attributes:

```
Attribute name: rmcap
Long form name: RTP-based media format capability
Type of attribute: session-level and media-level
Subject to charset: no
Purpose: associate RTP-based media capability number(s) with
media subtype and encoding parameters
Appropriate Values: see Section 3.3.1
Contact name: Flemming Andreasen, fandres@cisco.com

Attribute name: omcap
Long form name: non-RTP-based media format capability
Type of attribute: session-level and media-level
Subject to charset: no
Purpose: associate non-RTP-based media capability number(s) with
media subtype and encoding parameters
Appropriate Values: see Section 3.3.1
Contact name: Flemming Andreasen, fandreas@cisco.com
```

Attribute name: mfcap
Long form name: media format parameter capability
Type of attribute: session-level and media-level
Subject to charset: no
Purpose: associate media format attributes and
parameters with media format capabilities
Appropriate Values: see Section 3.3.2
Contact name: Flemming Andreasen, fandreas@cisco.com

Attribute name: mscap
Long form name: media-specific capability
Type of attribute: session-level and media-level
Subject to charset: no
Purpose: associate media-specific attributes and
parameters with media capabilities
Appropriate Values: see Section 3.3.3
Contact name: Flemming Andreasen, fandreas@cisco.com

Attribute name: lcfg
Long form name: latent configuration
Type of attribute: media-level
Subject to charset: no
Purpose: to announce supportable media streams
without offering them for immediate use.
Appropriate Values: see Section 3.3.5
Contact name: Flemming Andreasen, fandreas@cisco.com

Attribute name: sescap
Long form name: session capability
Type of attribute: session-level
Subject to charset: no
Purpose: to specify and prioritize acceptable
combinations of media stream configurations.
Appropriate Values: see Section 3.3.8
Contact name: Flemming Andreasen, fandreas@cisco.com

5.2. New SDP Capability Negotiation Option Tag

IANA has added the new option tag "med-v0", defined in this document, to the "SDP Capability Negotiation Option Capability Tags" registry created for RFC 5939 [RFC5939].

5.3. SDP Capability Negotiation Configuration Parameters Registry

IANA has changed the "SDP Capability Negotiation Potential Configuration Parameters" registry, currently registered and defined by RFC 5939 [RFC5939], as follows:

The name of the registry should be "SDP Capability Negotiation Configuration Parameters Registry" and it should contain a table with the following column headings:

- Encoding Name: The syntactical value used for the capability negotiation configuration parameter, as defined in RFC 5939 [RFC5939], Section 3.5.

- Descriptive Name: The name commonly used to refer to the capability negotiation configuration parameter.

- Potential Configuration Definition: A reference to the RFC that defines the configuration parameter in the context of a potential configuration attribute. If the configuration parameter is not defined for potential configurations, the string "N/A" (Not Applicable) MUST be present instead.

- Actual Configuration Definition: A reference to the RFC that defines the configuration parameter in the context of an actual configuration attribute. If the configuration parameter is not defined for actual configurations, the string "N/A" (Not Applicable) MUST be present instead.

- Latent Configuration Definition: A reference to the RFC that defines the configuration parameter in the context of a latent configuration attribute. If the configuration parameter is not defined for latent configurations, the string "N/A" (Not Applicable) MUST be present instead.

An IANA SDP Capability Negotiation Configuration registration MUST be documented in an RFC in accordance with the IETF Review policy [RFC5226]. Furthermore:

- The RFC MUST define the syntax and semantics of each new potential configuration parameter.

- The syntax MUST adhere to the syntax provided for extension configuration lists in RFC 5939 [RFC5939], Section 3.5.1, and the semantics MUST adhere to the semantics provided for extension configuration lists in RFC 5939 [RFC5939], Sections 3.5.1 and 3.5.2.

- Configuration parameters that apply to latent configurations MUST furthermore adhere to the syntax provided in Section 3.3.5 and the semantics defined overall in this document.

- Associated with each registration MUST be the encoding name for the parameter as well as a short descriptive name for it.

- Each registration MUST specify if it applies to
  - Potential configurations
  - Actual configurations
  - Latent configurations

5.4. SDP Capability Negotiation Configuration Parameter Registrations

IANA has registered the following capability negotiation configuration parameters:

```
Encoding Name: a
Descriptive Name: Attribute Configuration
Potential Configuration Definition: [RFC5939]
Actual Configuration Definition: [RFC5939]
Latent Configuration Definition: [RFC6871]

Encoding Name: t
Descriptive Name: Transport Protocol Configuration
Potential Configuration Definition: [RFC5939]
Actual Configuration Definition: [RFC5939]
Latent Configuration Definition: [RFC6871]

Encoding Name: m
Descriptive Name: Media Configuration
Potential Configuration Definition: [RFC6871]
Actual Configuration Definition: [RFC6871]
Latent Configuration Definition: [RFC6871]

Encoding Name: pt
Descriptive Name: Payload Type Number Mapping
Potential Configuration Definition: [RFC6871]
Actual Configuration Definition: [RFC6871]
Latent Configuration Definition: [RFC6871]

Encoding Name: mt
Descriptive Name: Media Type
Potential Configuration Definition: N/A
Actual Configuration Definition: N/A
Latent Configuration Definition: [RFC6871]
```

6. Security Considerations

The security considerations of RFC 5939 [RFC5939] apply for this document.

In RFC 5939 [RFC5939], it was noted that negotiation of transport protocols (e.g., secure and non-secure) and negotiation of keying methods and material are potential security issues that warrant integrity protection to remedy. Latent configuration support provides hints to the other side about capabilities supported for further offer/answer exchanges, including transport protocols and attribute capabilities, e.g., for keying methods. If an attacker can remove or alter latent configuration information to suggest that only non-secure or less-secure alternatives are supported, then he may be able to force negotiation of a less secure session than would otherwise have occurred. While the specific attack, as described here, differs from those described in RFC 5939 [RFC5939], the considerations and mitigation strategies are similar to those described in RFC 5939 [RFC5939].

Another variation on the above attack involves the session capability ("sescap") attribute defined in this document. The "sescap" enables a preference order to be specified for all the potential configurations, and that preference will take precedence over any preference indication provided in individual potential configuration attributes. Consequently, an attacker that can insert or modify a "sescap" attribute may be able to force negotiation of an insecure or less secure alternative than would otherwise have occurred. Again, the considerations and mitigation strategies are similar to those described in RFC 5939 [RFC5939].

The addition of negotiable media formats and their associated parameters, defined in this specification can cause problems for middleboxes that attempt to control bandwidth utilization, media flows, and/or processing resource consumption as part of network policy, but that do not understand the media capability negotiation feature. As for the initial SDP capability negotiation work [RFC5939], the SDP answer is formulated in such a way that it always carries the selected media encoding for every media stream selected. Pending an understanding of capabilities negotiation, the middlebox should examine the answer SDP to obtain the best picture of the media streams being established. As always, middleboxes can best do their job if they fully understand media capabilities negotiation.

7. Acknowledgements

This document is heavily influenced by the discussions and work done by the SDP Capability Negotiation design team. The following people in particular provided useful comments and suggestions to either the document itself or the overall direction of the solution defined herein: Cullen Jennings, Matt Lepinski, Joerg Ott, Colin Perkins, and Thomas Stach.

We thank Ingemar Johansson and Magnus Westerlund for examples that stimulated this work, and for critical reading of the document. We also thank Cullen Jennings, Christer Holmberg, and Miguel Garcia for their review of the document.

8. References

8.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC3264] Rosenberg, J. and H. Schulzrinne, "An Offer/Answer Model with Session Description Protocol (SDP)", RFC 3264, June 2002.

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

[RFC5226] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 5226, May 2008.

[RFC5234] Crocker, D. and P. Overell, "Augmented BNF for Syntax Specifications: ABNF", STD 68, RFC 5234, January 2008.

[RFC5939] Andreasen, F., "Session Description Protocol (SDP) Capability Negotiation", RFC 5939, September 2010.

8.2. Informative References

[RFC2198] Perkins, C., Kouvelas, I., Hodson, O., Hardman, V., Handley, M., Bolot, J., Vega-Garcia, A., and S. Fosse-Parisis, "RTP Payload for Redundant Audio Data", RFC 2198, September 1997.

[RFC4568] Andreasen, F., Baugher, M., and D. Wing, "Session Description Protocol (SDP) Security Descriptions for Media Streams", RFC 4568, July 2006.

[RFC4585] Ott, J., Wenger, S., Sato, N., Burmeister, C., and J. Rey, "Extended RTP Profile for Real-time Transport Control Protocol (RTCP)-Based Feedback (RTP/AVPF)", RFC 4585, July 2006.

[RFC4733] Schulzrinne, H. and T. Taylor, "RTP Payload for DTMF Digits, Telephony Tones, and Telephony Signals", RFC 4733, December 2006.

[RFC4867] Sjoberg, J., Westerlund, M., Lakaniemi, A., and Q. Xie, "RTP Payload Format and File Storage Format for the Adaptive Multi-Rate (AMR) and Adaptive Multi-Rate Wideband (AMR-WB) Audio Codecs", RFC 4867, April 2007.

[RFC5104] Wenger, S., Chandra, U., Westerlund, M., and B. Burman, "Codec Control Messages in the RTP Audio-Visual Profile with Feedback (AVPF)", RFC 5104, February 2008.

Authors' Addresses

Robert R Gilman
Independent
3243 W. 11th Ave. Dr.
Broomfield, CO 80020
USA

EMail: bob_gilman@comcast.net

Roni Even
Huawei Technologies
14 David Hamelech
Tel Aviv 64953
Israel

EMail: roni.even@mail01.huawei.com

Flemming Andreasen
Cisco Systems
Iselin, NJ
USA

EMail: fandreas@cisco.com