Internet Engineering Task Force (IETF)
Request for Comments: 6884
Category: Standards Track
ISSN: 2070-1721

Z. Fang
Qualcomm Incorporated

March 2013

RTP Payload Format
for the Enhanced Variable Rate Narrowband-Wideband Codec (EVRC-NW)

Abstract

This document specifies Real-time Transport Protocol (RTP) payload formats to be used for the Enhanced Variable Rate Narrowband-Wideband Codec (EVRC-NW). Three media type registrations are included for EVRC-NW RTP payload formats. In addition, a file format is specified for transport of EVRC-NW speech data in storage mode applications such as email.

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/rfc6884.

Copyright Notice

Table of Contents

1. Introduction

This document specifies the payload formats for packetization of EVRC-NW encoded speech signals into the Real-time Transport Protocol (RTP). It defines support for the header-free, interleaved/bundled, and compact bundle packet formats for the EVRC-NW codec as well as discontinuous transmission (DTX) support for EVRC-NW encoded speech transported via RTP. The EVRC-NW codec offers better speech quality than the EVRC and EVRC-B codecs and better capacity than the Enhanced Variable Rate Wideband Codec (EVRC-WB). EVRC-NW belongs to the EVRC family of codecs.

2. Conventions

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 [1].

3. Background

EVRC-NW is an extension of both the EVRC-B [2] and EVRC-WB [3] speech codecs developed in the Third Generation Partnership Project 2 (3GPP2) with support for DTX. It provides enhanced voice quality and high spectral efficiency.

The EVRC-NW codec operates on 20 ms frames, and the default sampling rate is 16 kHz (wideband). Input and output at the 8 kHz sampling rate (narrowband) is also supported. The EVRC-NW codec can operate in eight modes (0 to 7) as defined in 3GPP2 C.S0014-D [4]. EVRC-NW modes 0, 1, and 7 are interoperable with EVRC-WB. EVRC-NW modes 1 to 7 are interoperable with EVRC-B. EVRC-NW modes 0 to 6 use the full set or a subset of full rate, 1/2 rate, 1/4 rate, and 1/8 rate frames. EVRC-NW mode 7 uses only 1/2 rate and 1/8 rate frames. By default, EVRC-NW supports all narrowband modes (modes 1 to 7). The support of wideband mode (mode 0) is optional. Mode change among modes 1 to 7 (or among modes 0 to 7 if the receiver supports wideband mode) results in codec output bit-rate change but does not cause any decoding problems at the receiver. EVRC-NW provides a standardized solution for packetized voice applications that allow transitions between enhanced quality and increased capacity. The most important service addressed is IP telephony. Target devices can be IP phones or VoIP handsets, media gateways, voice messaging servers, etc.

4. EVRC-NW Codec

The EVRC-NW codec operates on 20 ms frames. It produces output frames of one of the four different sizes: 171 bits (Rate 1), 80 bits (Rate 1/2), 40 bits (Rate 1/4), or 16 bits (Rate 1/8). In addition, there are two zero-bit codec frame types: blank (null) frames and erasure frames. The default sampling rate is 16 kHz. Input and output at the 8 kHz sampling rate is also supported.

The frame type values and sizes of the associated codec data frames are listed in the table below:

Value	Rate	Total codec data frame size in bytes (and in bits)	
0	Blank (Null)	0	(0 bits)
1	1/8	2	(16 bits)
2	1/4	5	(40 bits)
3	1/2	10	(80 bits)
4	1	22	(171 bits; 5 bits padded at the end)
5	Erasure	0	(SHOULD NOT be transmitted by sender)

5. RTP Header Usage

The format of the RTP header is specified in RFC 3550 [5]. The EVRC-NW payload formats (Section 6) use the fields of the RTP header as specified in RFC 3550 [5].

EVRC-NW also has the capability to operate with 8 kHz sampled input/output signals. The decoder does not require a priori knowledge about the sampling rate of the original signal at the input of the encoder. The decoder output can be at 8 kHz or 16 kHz regardless of the sampling rate used at the encoder. Therefore, depending on the implementation and the electroacoustic audio capabilities of the devices, the input of the encoder and/or the output of the decoder can be configured at 8 kHz; however, a 16 kHz RTP clock rate MUST always be used. The RTP timestamp is increased by 320 for each 20 milliseconds.

The RTP header marker bit (M) SHALL be set to 1 if the first frame carried in the packet contains a speech frame that is the first in a talkspurt. For all other packets, the marker bit SHALL be set to zero (M=0).

6. Payload Format

Three RTP packet formats are supported for the EVRC-NW codec -- the interleaved/bundled packet format, the header-free packet format, and the compact bundled packet format. For all these formats, the operational details and capabilities of EVRC-NW, such as TOC, interleaving, DTX, and bundling, are exactly the same as those defined in EVRC [6], EVRC-B [2], and EVRC-WB [3], except that

1. the mode change request field in the interleaved/bundled packet format MUST be interpreted according to the definition of the RATE_REDUC parameter as described for EVRC-NW in 3GPP2 C.S0014-D [4].

2. the mode change request field in the interleaved/bundled packet format SHOULD be honored by an EVRC-NW encoding endpoint in a one-to-one session with a dedicated EVRC-NW decoding endpoint, such as in a two-party call or in a conference leg.

3. the reserved bit field in the first octet of the interleaved/bundled format has only one bit. Bit 1 of the first octet is an EVRC-NW wideband/narrowband encoding capability identification flag.

The media type audio/EVRCNW maps to the interleaved/bundled packet format, audio/EVRCNW0 maps to the header-free packet format, and audio/EVRCNW1 maps to the compact bundled packet format.

6.1. Encoding Capability Identification in EVRC-NW Interleaved/Bundled Format

The EVRC-NW interleaved/bundled format defines an encoding capability identification flag, which is used to signal the local EVRC-NW wideband/narrowband encoding capability at the time of construction of an RTP packet to the far end of a communication session. This capability identification flag allows the far end to use the MMM field in its outgoing (returning) EVRC-NW interleaved/bundled format packets to request the desired EVRC-NW wideband or narrowband encoding mode in accordance with the dynamic/instantaneous encoding capability information. See RFC 3558 [6] for the definition of the MMM field. The following examples illustrate a few scenarios where the encoding capability information is used:

- An end-to-end wideband communication is established first between two communication endpoints using the EVRC-NW interleaved/bundled format. The called endpoint becomes wideband encoding incapable during the call and makes the other end aware of this change by using the encoding capability identification flag. Based on the new information, the calling endpoint could change the MMM value in its outgoing EVRC-NW packets from mode 0 to mode 4 to request narrowband encoded traffic for bandwidth efficiency or from mode 0 to mode 1 for best perceptual quality.

- An end-to-end narrowband communication is established between a calling endpoint that is EVRC-NW wideband encoding capable and a called endpoint that is EVRC-NW wideband encoding incapable. The called endpoint becomes EVRC-NW wideband encoding capable during the call and makes the other end aware of this change using the encoding capability identification flag. Based on the new information, the calling endpoint could change the MMM value in its outgoing EVRC-NW packets from non-mode-0 to mode 0 to request wideband traffic.

The EVRC-NW interleaved/bundled format defines the encoding capability identification flag in bit 1 of the first octet, as illustrated in the figure below. The flag shall be set to zero (C=0) when the local EVRC-NW encoder is capable of mode 0 wideband encoding. The flag shall be set to one (C=1) when the local EVRC-NW encoder is capable of non-mode-0 narrowband encoding only. See RFC 3558 [6] for original definitions of other fields in the interleaved/bundled format.

```
     0                   1                   2                   3
     0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                         RTP Header                            |
    +=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+
    |R|C| LLL | NNN | MMM |  Count  |  TOC  |  ...  |  TOC  |padding|
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |        one or more codec data frames, one per TOC entry       |
    |                              ....                             |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Reserved (R): 1 bit

Reserved bit. MUST be set to zero by sender; SHOULD be ignored by receiver.

Encoding capability identification (C): 1 bit

Must be set to zero by sender to indicate wideband encoding capable or set to one to indicate narrowband encoding capable only.

C = 0 : mode 0 wideband encoding capable

= 1 : mode 0 wideband encoding incapable, i.e., narrowband encoding only.

7. Congestion Control Considerations

Congestion control for RTP is discussed in RFC 3550 [5] and in applicable RTP profiles, e.g., RFC3551 [7]. This document does not change those considerations.

Due to the header overhead, the number of frames encapsulated in each RTP packet influences the overall bandwidth of the RTP stream. Packing more frames in each RTP packet can reduce the number of packets sent and hence the header overhead, at the expense of increased delay and reduced error robustness.

8. Storage Format for the EVRC-NW Codec

The storage format is used for storing EVRC-NW encoded speech frames, e.g., as a file or email attachment.

The file begins with a magic number to identify the vocoder that is used. The magic number for EVRC-NW corresponds to the ASCII character string "#!EVRCNW\n", i.e., "0x23 0x21 0x45 0x56 0x52 0x43 0x4E 0x57 0x0A".

The codec data frames are stored in consecutive order, with a single TOC entry field, extended to one octet, prefixing each codec data frame. The TOC field is extended to one octet by setting the four most significant bits of the octet to zero. For example, a TOC value of 4 (a full-rate frame) is stored as 0x04. The Value column in the table in Section 4 provides the TOC values for corresponding frame types.

Speech frames lost in transmission and non-received frames MUST be stored as erasure frames (TOC value of 5) to maintain synchronization with the original media.

9. IANA Considerations

This document introduces a new EVRC-NW 'audio' media subtype.

9.1. Media Type Registrations

Following the guidelines in RFC 4855 [8] and RFC 6838 [9], this section registers new 'audio' media subtypes for EVRC-NW.

9.1.1. Registration of Media Type audio/EVRCNW

Type name: audio

Subtype name: EVRCNW

Required parameters: None

Optional parameters: These parameters apply to RTP transfer only.

mode-set-recv: A subset of EVRC-NW modes. Possible values are a comma-separated list of modes from the set {0,1,2,3,4,5,6,7} (see Table 2.6.1.2-1 in 3GPP2 C.S0014-D [4]). A decoder can use this attribute to inform an encoder of its preference to operate in a specified subset of modes. Absence of this parameter signals the mode set {1,2,3,4,5,6,7}.

ptime: See RFC 4566 [10].

maxptime: See RFC 4566.

maxinterleave: Maximum number for interleaving length (field LLL in the Interleaving Octet) [0..7]. The interleaving lengths used in the entire session MUST NOT exceed this maximum value. If not signaled, the maxinterleave length MUST be 5.

silencesupp: See Section 6.1 in RFC 4788.

dtxmax: See Section 6.1 in RFC 4788.

dtxmin: See Section 6.1 in RFC 4788.

hangover: See Section 6.1 in RFC 4788.

Encoding considerations:
This media type is framed binary data (see RFC 6838, Section 4.8) and is defined for transfer of EVRC-NW encoded data via RTP using the interleaved/bundled packet format specified in RFC 3558 [6].

Security considerations: See Section 16.

Interoperability considerations: None

Published specification:
The EVRC-NW vocoder is specified in 3GPP2 C.S0014-D [4]. The transfer method with the interleaved/bundled packet format via RTP is specified in RFC 3558 [6]. See Section 6 of RFC 6884 for details for EVRC-NW.

Applications that use this media type:
It is expected that many VoIP applications (as well as mobile applications) will use this type.

Additional information:

The following applies to stored-file transfer methods:

Magic number: #!EVRCNW\n (see Section 8)

File extensions: enw, ENW

Macintosh file type code: None

Object identifier or OID: None

EVRC-NW speech frames may also be stored in the file format "3g2" as defined in 3GPP2 C.S0050-B [14], which is identified using the media types "audio/3gpp2" or "video/3gpp2" registered by RFC 4393 [11].

Person & email address to contact for further information:
Zheng Fang <zfang@qualcomm.com>

Intended usage: COMMON

Restrictions on usage:
This media type can be used with the file format defined in Section 8 of RFC 6884 in contexts other than RTP. In the context of transfers over RTP, the RTP payload format specified in Section 4.1 of RFC 3558 [6] is used for this media type.

Author: Zheng Fang <zfang@qualcomm.com>

Change controller:
IETF Payload working group delegated from the IESG.

9.1.2. Registration of Media Type audio/EVRCNW0

Type name: audio

Subtype name: EVRCNW0

Required parameters: None

Optional parameters: These parameters apply to RTP transfer only.

mode-set-recv: A subset of EVRC-NW modes. Possible values are a comma-separated list of modes from the set {0,1,2,3,4,5,6,7} (see Table 2.6.1.2-1 in 3GPP2 C.S0014-D [4]). A decoder can use this attribute to inform an encoder of its preference to operate in a specified subset of modes. Absence of this parameter signals the mode set {1,2,3,4,5,6,7}.

ptime: See RFC 4566.

silencesupp: See Section 6.1 in RFC 4788.

dtxmax: See Section 6.1 in RFC 4788.

dtxmin: See Section 6.1 in RFC 4788.

hangover: See Section 6.1 in RFC 4788.

Encoding considerations:
This media type is framed binary data (see RFC 6838, Section 4.8) and is defined for transfer of EVRC-NW encoded data via RTP using the header-free packet format specified in RFC 3558 [6].

Security considerations: See Section 16.

Interoperability considerations: None

Published specification:
The EVRC-NW vocoder is specified in 3GPP2 C.S0014-D [4]. The transfer method with the header-free packet format via RTP is specified in RFC 3558 [6].

Applications that use this media type:
It is expected that many VoIP applications (as well as mobile applications) will use this type.

Additional information: None

Person & email address to contact for further information:
Zheng Fang <zfang@qualcomm.com>

Intended usage: COMMON

Restrictions on usage:
This media type depends on RTP framing and hence is only defined for transfer via RTP [5]. The RTP payload format specified in Section 4.2 of RFC 3558 [6] SHALL be used. This media type SHALL NOT be used for storage or file transfer; instead, audio/EVRCNW SHALL be used.

Author: Zheng Fang <zfang@qualcomm.com>

Change controller:
IETF Payload working group delegated from the IESG.

9.1.3. Registration of Media Type audio/EVRCNW1

Type name: audio

Subtype name: EVRCNW1

Required parameters: None

Optional parameters: These parameters apply to RTP transfer only.

mode-set-recv: A subset of EVRC-NW modes. Possible values are a comma-separated list of modes from the set {0,1} (see Table 2.6.1.2-1 in 3GPP2 C.S0014-D [4]). A decoder can use this attribute to inform an encoder of its preference to operate in a specified subset of modes. A value of 0 signals support for wideband fixed rate (full or half rate, depending on the value of the ’fixedrate’ parameter). A value of 1 signals narrowband fixed rate (full or half rate, depending on the value of the ’fixedrate’ parameter). Absence of this parameter signals mode 1.

ptime: See RFC 4566.

maxptime: See RFC 4566.

fixedrate: Indicates the EVRC-NW rate of the session while in single rate operation. Valid values include 0.5 and 1, where a value of 0.5 indicates the 1/2 rate while a value of 1 indicates the full rate. If this parameter is not present, 1/2 rate is assumed.

silencesupp: See Section 6.1 in RFC 4788.

dtxmax: See Section 6.1 in RFC 4788.

dtxmin: See Section 6.1 in RFC 4788.

hangover: See Section 6.1 in RFC 4788.

Encoding considerations:
This media type is framed binary data (see RFC 6838, Section 4.8) and is defined for transfer of EVRC-NW encoded data via RTP using the compact bundled packet format specified in RFC 4788.

Security considerations: See Section 16.

Interoperability considerations: None

Published specification:
The EVRC-NW vocoder is specified in 3GPP2 C.S0014-D [4]. The transfer method with the compact bundled packet format via RTP is specified in RFC 4788.

Applications that use this media type:
It is expected that many VoIP applications (as well as mobile applications) will use this type.

Additional information: None

Person & email address to contact for further information:
Zheng Fang <zfang@qualcomm.com>

Intended usage: COMMON

Restrictions on usage:
This media type depends on RTP framing and hence is only defined for transfer via RTP [5]. The RTP payload format specified in Section 4 of RFC 4788 SHALL be used. This media type SHALL NOT be used for storage or file transfer; instead, audio/EVRCNW SHALL be used.

Author: Zheng Fang <zfang@qualcomm.com>

Change controller:
IETF Payload working group delegated from the IESG.

10. SDP Mode Attributes for EVRC-NW

'mode-set-recv' can be used by a decoder to inform an encoder of its preference to operate in a specified subset of modes. Note that indicating a preference implicitly indicates support for that capability. If mode 0 is not preferred for media type EVRCNW0 or EVRCNW1, then there is no indication that mode 0 is supported. However, absence of this parameter or absence of mode 0 in this parameter for media type EVRCNW shall not preclude mode 0 support during a call where mode 0 may be requested via the MMM field.

1. To inform other nodes of its capability for wideband mode support: a decoder can always decode all the narrowband modes (modes 1 to 7). Unless the decoder indicates support of mode 0 (i.e., preference) in this parameter or in the MMM mode request field in the interleaved/bundled payload format, an encoder at the other side shall not operate in mode 0.

2. To indicate a preference to operate in a subset of modes: a set has been defined so that several modes can be expressed as a preference in one attempt. For instance, the set {4,5,6,7} signals that the receiver prefers that the sender operate in bandwidth-efficient narrowband modes of EVRC-NW.

Note that during an active call session using the interleaved/bundled packet format, the MMM mode request received from a communication partner can contain a mode request different than the values in the last mode-set-recv attribute. The partner's EVRC-NW wideband decoding capability is determined by the latest mode-set-recv attribute or MMM mode request field. For example, a mode request with MMM=0 from a communication partner is an implicit indication of the partner's EVRC-NW wideband decoding capability and preference. An EVRC-NW wideband-capable node receiving the request can operate in wideband mode. A mode request with MMM=1, 2, ..., or 7 from a communication partner is an implicit indication of the partner's

EVRC-NW narrowband decoding preference. The encoder of an EVRC-NW node receiving the request shall honor the request and operate in narrowband mode.

'sendmode' is used as a Session Description Protocol (SDP) mode attribute in EVRC [6], EVRC-B [2], and EVRC-WB [3]. However, it is deprecated in EVRC-NW.

11. Mode Change Request/Response Considerations

The interleaved/bundled packet format for the EVRC family of vocoders supports a 3-bit field (MMM) that a communication node can use to indicate its preferred compression mode to an opposite node. The concept of the compression mode (also known as Capacity Operating Point) was introduced to allow a controlled trade-off between voice quality and channel capacity. The notion makes it possible to exercise vocoders at the highest possible (average) bit-rate (hence, highest voice quality) when the network is lightly loaded. Conversely, once the network load increases, the vocoders can be requested to operate at lower average bit-rates so as to absorb the additional network load without causing an undue increase in the frame-erasure rates; the underlying premise is that while a higher bit-rate improves vocoder performance, it also increases the network load, risking a sharp decline in voice quality should the frame-erasure rate be too high. By contrast, a lower bit-rate mode of operation can result in accommodation of the additional network load without causing unduly high frame-erasure rates, resulting in better overall quality despite the inherently lower voice quality of the lower bit-rate mode of the vocoder.

Accordingly, the MMM field should be used to request the far end to transmit compressed speech using a mode that provides the best balance between voice quality and capacity. However, in the case of mobile-mobile calls, for example, there are two wireless sides involved, each with a potentially different network load level and hence a different preferred mode. In such cases, achieving optimal end-to-end performance depends on coherent management of the operative mode by the two sides. This requires that even if the local node prefers a higher bit-rate vocoder mode, it should adjust to a lower bit-rate mode if requested by the far end, in order to avoid potentially high frame-erasure rates due to heavy load at the far-end network. For similar reasons, in cases where a mode requested by the far end should not be supported, it might still be beneficial to consider switching to a supported vocoder mode corresponding to a lower average bit-rate than requested. It is recommended that the next lower average bit-rate supported vocoder mode be used for encoding when a mode requested by the far end is not supported.

A wideband-capable endpoint can use the information conveyed by the C-bit of the RTP payload header to determine the optimal mode to request of the far end. If the far end cannot provide mode 0 packets (C-bit=1), then the choice of MMM can be based strictly on the local network load. If the C-bit indicates the remote end’s mode 0 encoding capability (C-bit=0), then even if the local network load is not light, mode 0 can be requested knowing definitively that it will be supported. This will permit operators to treat wideband-capable mobiles preferentially, should they wish to adopt such policy.

12. Mapping EVRC-NW Media Type Parameters into SDP

Information carried in the media type specification has a specific mapping to fields in the Session Description Protocol (SDP) [10], which is commonly used to describe RTP sessions. When SDP is used to specify sessions employing EVRC-NW encoded speech, the mapping is as follows.

- The media type ("audio") goes in SDP "m=" as the media name.
- The media subtype ("EVRCNW", "EVRCNW0", or "EVRCNW1") goes in SDP "a=rtpmap" as the encoding name.
- The optional parameters ’ptime and ’maxptime’ (for subtypes EVRCNW and EVRCNW1) go in the SDP "a=ptime" and "a=maxptime" attributes, respectively.
- Any remaining parameters (for subtypes EVRCNW, EVRCNW0, and EVRCNW1) go in the SDP "a=fmtp" attribute by copying them from the media type string as a semicolon-separated list of parameter=value pairs.

13. Offer-Answer Model Considerations for EVRC-NW

The following considerations apply when using the SDP offer-answer procedures of RFC 3264 [12] to negotiate the use of EVRC-NW payload in RTP:

- Since EVRC-NW is an extension of both EVRC-B and EVRC-WB, the offerer SHOULD also announce EVRC-B and EVRC-WB support in its "m=audio" lines, with EVRC-NW as the preferred codec. This will allow interoperability with an answerer that supports only EVRC-B and/or EVRC-WB.

Below is an example of such an offer:

```
m=audio 55954 RTP/AVP 98 99 100
a=rtpmap:98 EVRCNW0/16000
a=rtpmap:99 EVRCWB0/16000
a=rtpmap:100 EVRCB0/8000
a=fmtp:98 mode-set-recv=0,1,2,3,4,5,6
a=fmtp:99 mode-set-recv=0,4
a=fmtp:100 recvmode=0
```

If the answerer supports EVRC-NW, then the answerer can keep the payload type 98 in its answer and the conversation can be done using EVRC-NW. Otherwise, if the answerer supports only EVRC-WB and/or EVRC-B, then the answerer will leave only the payload type 99 and/or 100, respectively, in its answer and the conversation will be done using EVRC-WB and/or EVRC-B, respectively.

An example answer for the above offer:

```
m=audio 55954 RTP/AVP 98
a=rtpmap:98 EVRCNW0/16000
a=fmtp:98 mode-set-recv=4
```

- ’mode-set-recv’ is a unidirectional receive-only parameter.

- An offerer can use ’mode-set-recv’ to request that the remote sender’s encoder be limited to the list of modes signaled in ’mode-set-recv’. A remote sender MAY ignore ’mode-set-recv’ requests. However, a remote sender shall not assume the other side can support mode 0, unless the offer includes mode 0 explicitly in ’mode-set-recv’ or the remote sender receives mode requests with MMM=0 from the communication partner during an active call using the EVRC-NW interleaved/bundled format.

- The parameters ’maxptime’ and ’ptime’ will in most cases not affect interoperability; however, the setting of the parameters can affect the performance of the application. The SDP offer-answer handling of the ’ptime’ parameter is described in RFC 3264 [12]. The ’maxptime’ parameter MUST be handled in the same way.

- For a sendonly stream, the ’mode-set-recv’ parameter is not useful and SHOULD NOT be used.

- When using EVRCNW1, the entire session MUST use the same fixed rate and mode (0-Wideband or 1-Narrowband).

- For additional rules that MUST be followed while negotiating DTX parameters, see Section 6.8 in RFC 4788 [2].

- Any unknown parameter in an SDP offer MUST be ignored by the receiver and MUST NOT be included in the SDP answer.

14. Declarative SDP Considerations

For declarative use of SDP in the Session Announcement Protocol (SAP) [15] and the Real Time Streaming Protocol (RTSP) [16], the following considerations apply:

- Any 'maxptime' and 'ptime' values should be selected with care to ensure that the session's participants can achieve reasonable performance.

- The payload format configuration parameters are all declarative, and a participant MUST use the configuration(s) that is provided for the session. More than one configuration MAY be provided if necessary by declaring multiple RTP payload types; however, the number of types SHOULD be kept small. For declarative examples, see Section 15.

- The usage of unidirectional receive-only parameters, such as 'mode-set-recv', should be excluded in any declarations, since these parameters are meaningless in one-way streaming applications.

15. Examples

Some example SDP session descriptions utilizing EVRC-NW encodings follow. In these examples, long a=fmtp lines are folded to meet the column width constraints of this document. The backslash ("\") at the end of a line and the carriage return that follows it should be ignored. Note that media subtype names are case-insensitive. Parameter names are case-insensitive both in media types and in the mapping to the SDP a=fmtp attribute.

Example usage of EVRCNW if wideband mode is supported:

```
m=audio 49120 RTP/AVP 97 98 99
a=rtpmap:97 EVRCNW/16000
a=rtpmap:98 EVRCWB/16000
a=rtpmap:99 EVRCB/8000
a=fmtp:97 mode-set-recv=0,1,2,3,4,5,6
a=fmtp:98 mode-set-recv=0,4
a=fmtp:99 recvmode=0
a=maxptime:120
```

Example usage of EVRCNW if wideband mode is not supported:

```
m=audio 49120 RTP/AVP 97 98 99
a=rtpmap:97 EVRCNW/16000
a=rtpmap:98 EVRCWB/16000
a=rtpmap:99 EVRCB/8000
a=fmtp:97 mode-set-recv=1,2,3,4,5,6
a=fmtp:98 mode-set-recv=4
a=fmtp:99 recvmode=0
a=maxptime:120
```

Example usage of EVRCNW0:

```
m=audio 49120 RTP/AVP 97 98 99
a=rtpmap:97 EVRCNW0/16000
a=rtpmap:98 EVRCWB0/16000
a=rtpmap:99 EVRCB0/8000
a=fmtp:97 mode-set-recv=0,1,2,3,4,5,6
a=fmtp:98 mode-set-recv=0,4
a=fmtp:99 recvmode=0
```

Example SDP answer from a media gateway requesting a terminal to limit its encoder operation to EVRC-NW mode 4.

```
m=audio 49120 RTP/AVP 97
a=rtpmap:97 EVRCNW0/16000
a=fmtp:97 mode-set-recv=4
```

Example usage of EVRCNW1:

```
m=audio 49120 RTP/AVP 97 98 99
a=rtpmap:97 EVRCNW1/16000
a=rtpmap:98 EVRCWB1/16000
a=rtpmap:99 EVRCB1/8000
a=fmtp:97 fixedrate=0.5
a=fmtp:98 fixedrate=0.5
a=fmtp:99 fixedrate=0.5
a=maxptime:100
```

Example usage of EVRCNW with DTX with silencesupp=1:

```
m=audio 49120 RTP/AVP 97 98 99
a=rtpmap:97 EVRCNW/16000
a=rtpmap:98 EVRCWB/16000
a=rtpmap:99 EVRCB/8000
a=fmtp:97 silencesupp=1;dtxmax=32;dtxmin=12;hangover=1; \
mode-set-recv=0,1,2,3,4,5,6
a=fmtp:98 silencesupp=1;dtxmax=32;dtxmin=12;hangover=1; \
mode-set-recv=0,4
a=fmtp:99 recvmode=0
a=maxptime:120
```

Example usage of EVRCNW with DTX with silencesupp=0:

```
m=audio 49120 RTP/AVP 97 98 99
a=rtpmap:97 EVRCNW/16000
a=rtpmap:98 EVRCWB/16000
a=rtpmap:99 EVRCB/8000
a=fmtp:97 silencesupp=0;dtxmax=32;dtxmin=12;hangover=1; \
mode-set-recv=0,1,2,3,4,5,6
a=fmtp:98 silencesupp=0;dtxmax=32;dtxmin=12;hangover=1; \
mode-set-recv=0,4
a=fmtp:99 recvmode=0
a=maxptime:120
```

Example offer-answer exchange between EVRC-NW and legacy EVRC-B (RFC 4788):

Offer:

```
m=audio 55954 RTP/AVP 97 98 99
a=rtpmap:97 EVRCNW0/16000
a=rtpmap:98 EVRCWB0/16000
a=rtpmap:99 EVRCB0/8000
a=rtpmap:97 mode-set-recv=0,1,2,3,4,5,6
a=fmtp:98 mode-set-recv=0,4
a=fmtp:99 recvmode=0
```

Answer:

```
m=audio 55954 RTP/AVP 99
a=rtpmap:99 EVRCB0/8000
```

Example offer-answer exchange between EVRC-NW and legacy EVRC-WB (RFC 5188):

Offer:

```
m=audio 55954 RTP/AVP 97 98 99
a=rtpmap:97 EVRCNW0/16000
a=rtpmap:98 EVRCWB0/16000
a=rtpmap:99 EVRCB0/8000
a=rtpmap:97 mode-set-recv=0,1,2,3,4,5,6
a=fmtp:98 mode-set-recv=0,4
a=fmtp:99 recvmode=0
```

Answer:

```
m=audio 55954 RTP/AVP 98 99
a=rtpmap:98 EVRCWB0/16000
```

16. Security Considerations

Since compression is applied to the payload formats end-to-end, and the encodings do not exhibit significant non-uniformity, implementations of this specification are subject to all the security considerations specified in RFC 3558 [6]. Implementations using the payload defined in this specification are subject to the security considerations discussed in RFC 3558 [6], RFC 3550 [5], and any appropriate profile (for example, RFC 3551 [7]). Additional security considerations are described in RFC 6562 [13].

17. References

17.1. Normative References

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

[2] Xie, Q. and R. Kapoor, "Enhancements to RTP Payload Formats for EVRC Family Codecs", RFC 4788, January 2007.

[3] Desineni, H. and Q. Xie, "RTP Payload Format for the Enhanced Variable Rate Wideband Codec (EVRC-WB) and the Media Subtype Updates for EVRC-B Codec", RFC 5188, February 2008.

[4] "Enhanced Variable Rate Codec, Speech Service Options 3, 68, 70, and 73 for Wideband Spread Spectrum Digital Systems", 3GPP2 C.S0014-D v3.0, October 2010, <http://www.3gpp2.org/public_html/specs/C.S0014-D_v3.0_EVRC.pdf>.

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

[6] Li, A., "RTP Payload Format for Enhanced Variable Rate Codecs (EVRC) and Selectable Mode Vocoders (SMV)", RFC 3558, July 2003.

[7] Schulzrinne, H. and S. Casner, "RTP Profile for Audio and Video Conferences with Minimal Control", STD 65, RFC 3551, July 2003.

[8] Casner, S., "Media Type Registration of RTP Payload Formats", RFC 4855, February 2007.

[9] Freed, N., Klensin, J., and T. Hansen, "Media Type Specifications and Registration Procedures", BCP 13, RFC 6838, January 2013.

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

[11] Garudadri, H., "MIME Type Registrations for 3GPP2 Multimedia Files", RFC 4393, March 2006.

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

[13] Perkins, C. and JM. Valin, "Guidelines for the Use of Variable Bit Rate Audio with Secure RTP", RFC 6562, March 2012.

17.2. Informative References

[14] "3GPP2 File Formats for Multimedia Services", 3GPP2 C.S0050-B v1.0, May 2007, <http://www.3gpp2.org/public_html/specs/C.S0050-B_v1.0_070521.pdf>.

[15] Handley, M., Perkins, C., and E. Whelan, "Session Announcement Protocol", RFC 2974, October 2000.

[16] Schulzrinne, H., Rao, A., and R. Lanphier, "Real Time Streaming Protocol (RTSP)", RFC 2326, April 1998.

Author's Address

Zheng Fang
Qualcomm Incorporated
5775 Morehouse Drive
San Diego, CA 92126
USA

Phone: +1 858 651 9484
EMail: zfang@qualcomm.com
URI: http://www.qualcomm.com