              Enhancements to RTP Payload Formats for EVRC Family Codecs

Status of This Memo

Copyright Notice

Abstract

   This document updates the Enhanced Variable Rate Codec (EVRC) RTP
   payload formats defined in RFC 3558 with several enhancements and
   extensions. In particular, it defines support for the header-free
   and interleaved/bundled packet formats for the EVRC-B codec, a new
   compact bundled format for the EVRC and EVRC-B codecs, as well as
   discontinuous transmission (DTX) support for EVRC and EVRC-B-encoded
   speech transported via RTP. Voice over IP (VoIP) applications
   operating over low bandwidth dial-up and wireless networks require
   such enhancements for efficient use of the bandwidth.

Table of Contents

1. Introduction

   This document defines support for the header-free and interleaved/
   bundled packet formats for the EVRC-B codec, a new compact bundled
   format for the EVRC and EVRC-B codecs, as well as discontinuous
   transmission (DTX) support for EVRC and EVRC-B-encoded speech
   transported via RTP. Voice over IP (VoIP) applications operating
   over low bandwidth dial-up and wireless networks require such EVRC
   RTP payload capabilities for efficient use of the bandwidth.

1.1. Support of EVRC-B Codec

   EVRC-B [3] is an extension to EVRC [2] developed in the Third
   Generation Partnership Project 2 (3GPP2). EVRC-B [3] compresses each
   20 milliseconds of 8000Hz, 16-bit sampled speech input into output
   frames of one of the four different sizes: Rate 1 (171 bits), Rate
   1/2 (80 bits), Rate 1/4 (40 bits), or Rate 1/8 (16 bits). In
   addition, there are two zero-bit codec frame types: null frames and
   erasure frames, similar to EVRC [2]. One significant enhancement in
   EVRC-B is the use of 1/4-rate frames that were not used in EVRC.
   This provides lower average data rates (ADRs) compared to EVRC, for a
   given voice quality.

   Since speech frames encoded by EVRC-B are different from those
   encoded by EVRC, EVRC-B and EVRC codecs do not interoperate with each
   other. At the initiation of an RTP session, the RTP sender and
   receiver need to indicate (e.g., using MIME subtypes that are
   separate from those of EVRC) that EVRC-B is to be used for the
   ensuing session.

1.2. Compact (Header-free) Bundled Format

   The current interleaved/bundled packet format defined in RFC 3558
   allows bundling of multiple speech frames of different rates in a
   single RTP packet, sending mode change requests, and interleaving.
   To support these functions, a Table of Contents (ToC) is used in each
   RTP packet, in addition to the standard RTP header. The size of the
   ToC varies depending on the number of EVRC frames carried in the
   packet [4].

   The current header-free packet format defined in RFC 3558 is more
   compact and optimized for use over wireless links. It eliminates the
   need for a ToC by requiring that each RTP packet contain only one
   speech frame (of any allowable rate), i.e., bundling is not allowed.
   Moreover, interleaving and mode change requests are not supported in
   the header-free format [4].

The compact bundled format described in this document presents the
user an alternative to the header-free format defined in RFC 3558.
This format allows bundling of multiple EVRC or EVRC-B frames without
the addition of extra headers, as would be in the case of the
interleaved/bundled format. However, in order to use this compact
bundled format, only one EVRC/EVRC-B rate (full rate or 1/2 rate) can
be used in the session. Similar to the header-free format defined in
RFC 3558, interleaving and mode change requests are not supported in
the compact bundled format.

1.3. Discontinuous Transmission (DTX)

Information carried in frames of EVRC and EVRC-B codecs varies little
during periods of silence. The transmission of these frames across
the radio interface in a wireless system is expensive, in terms of
capacity; therefore, suppression of these frames is desirable. Such
an operation is called DTX, also known as silence suppression.

In general, when DTX/silence suppression is applied, the first few
frames of silence may be transmitted at the beginning of the period
of silence to establish background noise. Then, a portion of the
stream of subsequent silence frames is not transmitted, and is
discarded at the sender. At the receiver, background or comfort
noise may be generated by using the previously received silence
frames.

The full detail of DTX/silence suppression operation can be found in
DTX [8] as well as in RFC 3551 [9], and in RFC 3558 [4]. This
document only defines the additional optional MIME parameters
(silencesupp, dtxmax, dtxmin, and hangover) for setting up a DTX/
silence suppression session, where "silencesupp" is for indicating
the capability and willingness of using DTX/silence suppression;
"dtxmax" and "dtxmin", for indicating the desired range of DTX update
interval; and "hangover", for indicating the desired number of
silence frames at the beginning of each silence period to establish
background noise at the receiver (see Section 6.1 for detailed
definition).

The EVRC and EVRC-B codecs, in variable-rate operation mode, send
1/8-rate frames during periods of silence, while in single-rate
operation mode (see Section 4), silence is encoded and sent in frames
of the same rate as that of speech frames. The DTX parameters
defined in this document apply to 1/8-rate frames in the variable-
rate mode and to silence frames in the single-rate operation mode.

For simplicity, in the rest of this document the term "silence frame"
refers either to an 1/8-rate frame in variable-rate operation or a
frame that contains only silence in the signal-rate operation.
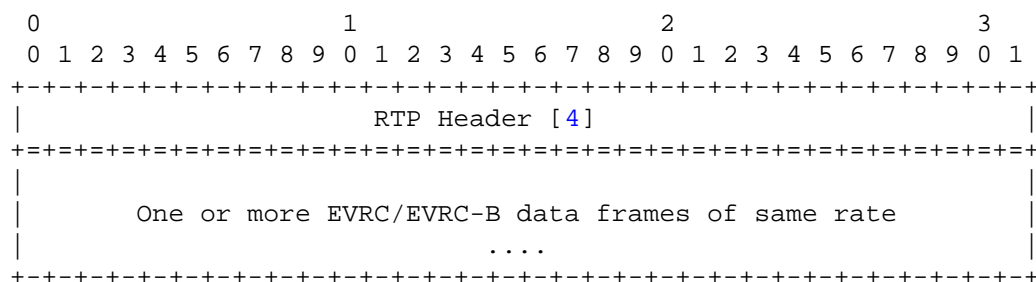
2. Conventions

   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. EVRC-B Codec

   Three RTP packet formats are supported for the EVRC-B codec: the
   interleaved/bundled packet format, the header-free packet format, and
   the compact bundled packet format. For the interleaved/bundled and
   header-free packet formats, the operational details and capabilities,
   such as ToC, interleaving, and bundling, of EVRC-B, are exactly the
   same as those of EVRC, as defined in RFC 3558 [4], except that the
   mode change request field in the ToC MUST be interpreted according to
   the definition of the RATE_REDUC parameter in EVRC-B [3]. The
   compact bundled packet format for EVRC-B is defined in Section 4 of
   this document.

4. Compact Bundled Format

   A packet in the compact bundled format consists of an RTP header,
   followed by a sequence of one or more consecutive EVRC/EVRC-B codec
   data frames of the same rate, as shown below:

```
    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 [4]                         |
   +=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+=+
   |                                                              |
   |       One or more EVRC/EVRC-B data frames of same rate       |
   |                           ....                               |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

   The codec data frames MUST be generated from the output of the codec
   following the procedure described in Section 5.2 in RFC 3558 [4], and
   all MUST be of the same rate and size.

4.1. Single-Rate Operation

   As mentioned earlier, in order to use the compact bundled format, all
   the EVRC/EVRC-B data frames in the session MUST be of the same rate.
   This packet format may carry only full or half-rate frames.

For a session that uses the compact bundled format, the rate for the
session can be determined during the session setup signaling, for
example, via Session Description Protocol (SDP) exchanges. See
Section 6 below for more details.

5. Storage Format for EVRC-B Codec

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

The file begins with a magic number to identify the vocoder that is
used. The magic number for EVRC-B corresponds to the ASCII character
string:

        "#!EVRC-B\n"
        (or 0x2321 0x4556 0x5243 0x2d42 0x0a in hexadecimal).

Note that the "\n" is an important part of both this magic number and
the "#!EVRC\n" magic number defined in Section 11 of RFC 3558, and
the "\n" MUST be included in any comparison of either magic number,
since, otherwise, a prefix of the EVRC-B magic number could be
mistaken for the EVRC magic number.

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, as defined in Section 5.1 of [4], 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.

Speech frames lost in transmission and non-received frames MUST be
stored as erasure frames to maintain synchronization with the
original media.

6. Media Type Definitions

6.1. Registration of Media Type EVRC1

Type name: audio

Subtype names: EVRC1

Required parameters: none

   Optional parameters:

      ptime: See RFC 4566 [7].

      maxptime: The maximum amount of media that can be encapsulated in
         each packet, expressed as time in milliseconds. The time MUST
         be calculated as the sum of the time the media present in the
         packet represents. The time SHOULD be a multiple of the
         duration of a single codec data frame (20 msec). If not
         signaled, the default maxptime value MUST be 200 milliseconds.

      fixedrate: Indicates the EVRC 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: Permissible values are 0 and 1. A value of 1
         indicates that the sender of this parameter: a) is capable of
         receiving silence-suppressed speech using DTX, AND b) is
         capable of and will send out silence-suppressed speech using
         DTX, unless the other end indicates that it does not want to
         receive silence-suppressed speech using DTX.

         A value of 0 indicates that the sender of this parameter: a)
         does NOT want to receive silence-suppressed speech using DTX,
         AND b) will NOT send out silence-suppressed speech using DTX.

         If this parameter is not present, the default value 1 MUST be
         assumed. If the RTP receiver indicates through the use of SIP
         signaling or other means that it is incapable of or unwilling
         to use silence suppression using DTX, silence suppression using
         DTX as specified in this document MUST NOT be used for the
         session.

      dtxmax: Permissible values are from 0 to 255. Indicates the
         maximum DTX update interval in number of frames. During DTX,
         the RTP sender occasionally updates the RTP receiver about the
         change in background noise characteristics, etc., by sending a
         new silence frame to the RTP receiver. The RTP receiver may
         use 'dtxmax' to indicate to the RTP sender the maximum interval
         (in number of frames) between any two DTX updates it expects to
         receive from the RTP sender.

         If this parameter is not present in a session that uses DTX,
         the default value 32, as specified in [8], MUST be assumed.
         This parameter MUST be ignored if silence suppression using DTX
         is not used for the session.

Note also that if the RTP receiver elects to detect DTX using
dtxmax, the dtxmax parameter will affect the amount of delay
the RTP receiver sees before detecting DTX in the stream.

dtxmin: Permissible values are from 0 to 255. Indicates the
minimum DTX update interval in number of frames. The RTP
receiver may use 'dtxmin' to indicate to the RTP sender the
minimal interval (in number of frames) between any two DTX
updates it expects to receive from the RTP sender.

If this parameter is not present, the default value 12, as
specified in [8] MUST be assumed. This parameter MUST be
ignored if silence suppression using DTX is not used for the
session.

hangover: Permissible values are from 0 to 255. Indicates the
number of consecutive silence frames transmitted at the end of
an active speech interval but before the DTX interval begins.
When setting up an RTP session that uses DTX, an RTP receiver
can use this parameter to signal the number of silence frames
it expects to receive before the beginning of DTX. While
hangover=0 is allowed, it is RECOMMENDED that hangover be set
to 1 or greater since the presence of silence frames at the end
of an active speech can help the RTP receiver to identify the
beginning of the DTX period.

If this parameter is not present for a session that uses DTX,
the default value 1, as specified in [8] MUST be assumed. This
parameter MUST be ignored if silence suppression using DTX is
not used for the session.

Encoding considerations:
   This media type is framed binary data (see RFC 4288, Section 4.8)
   and is defined for transfer of EVRC-encoded data via RTP, using
   the compact bundled format as described in RFC 4788.

Security considerations: See Section 9 of RFC 4788.

Interoperability considerations: none

Published specification:
   The EVRC vocoder is specified in 3GPP2 C.S0014 [2]. Transfer
   method with compact bundled RTP format 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:
      Qiaobing Xie <Qiaobing.Xie@motorola.com>

   Intended usage: COMMON

   Restrictions on usage:
      This media type depends on RTP framing; hence, it is only defined
      for transfer via RTP (RFC 3550 [5]). Transfer within other
      framing protocols is not defined at this time.

   Author:
      Qiaobing Xie

   Change controller:
      IETF Audio/Video Transport working group delegated from the IESG.

6.2. Registration of Media Type EVRCB

   Type name: audio

   Subtype names: EVRCB

   Required parameters: none

   Optional parameters:

      ptime: see RFC 4566 [7].

      maxptime: The maximum amount of media that can be encapsulated in
         each packet, expressed as time in milliseconds. The time MUST
         be calculated as the sum of the time the media present in the
         packet represents. The time SHOULD be a multiple of the
         duration of a single codec data frame (20 msec). If not
         signaled, the default maxptime value MUST be 200 milliseconds.

      maxinterleave: Maximum number for interleaving length (field LLL
         in the Interleaving Octet). 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 for definition. If this parameter
         is not present, the default value 1 MUST be assumed.

   dtxmax: see Section 6.1

   dtxmin: see Section 6.1

   hangover: see Section 6.1

Encoding considerations:
   This media type is framed binary data (see RFC 4288, Section 4.8)
   and is defined for transfer of EVRC-B-encoded data via RTP using
   the Interleaved/Bundled packet format specified in RFC 3558 [4].

Security considerations: See Section 9 of RFC 4788.

Interoperability considerations: none

Published specification:
   The EVRC-B vocoder is specified in 3GPP2 C.S0014-B [3]. Transfer
   method with Interleaved/Bundled packet format via RTP is specified
   in RFC 3558.

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 information applies for storage format only.

   Magic number: #!EVRC-B\n (see Section 5 of RFC 4788)
   File extensions: evb, EVB
   Macintosh file type code: None
   Object identifier or OID: None

Person & email address to contact for further information:
   Qiaobing Xie <Qiaobing.Xie@motorola.com>

Intended usage: COMMON

Restrictions on usage:
   This media type may be used with RTP framing (RFC 3550 [5]) and as
   a storage format. When used with RTP, the procedures in Section 3
   MUST be followed. In all other contexts, the storage format
   defined in Section 5 MUST be used.

Author:
   Qiaobing Xie

   Change controller:
      IETF Audio/Video Transport working group delegated from the IESG.

6.3. Registration of Media Type EVRCB0

   Type name: audio

   Subtype names: EVRCB0

   Required parameters: none

   Optional parameters:

      silencesupp: see Section 6.1 for definition. If this parameter
         is not present, the default value 1 MUST be assumed.

      dtxmax: see Section 6.1

      dtxmin: see Section 6.1

      hangover: see Section 6.1

   Encoding considerations:
      This media type is framed binary data (see RFC 4288, Section 4.8)
      and is defined for transfer of EVRC-B-encoded data via RTP using
      the Header-Free packet format specified in RFC 3558 [4].

   Security considerations: See Section 9 of RFC 4788.

   Interoperability considerations: none

   Published specification:
      The EVRC-B vocoder is specified in 3GPP2 C.S0014-B [3]. Transfer
      method with Header-Free packet format via RTP is specified in RFC
      3558 and 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:
      Qiaobing Xie <Qiaobing.Xie@motorola.com>

   Intended usage: COMMON

   Restrictions on usage:
      This media type depends on RTP framing; hence, it is only defined
      for transfer via RTP (RFC 3550 [5]). Transfer within other
      framing protocols is not defined at this time.

   Author:
      Qiaobing Xie

   Change controller:
      IETF Audio/Video Transport working group delegated from the IESG.

6.4. Registration of Media Type EVRCB1

   Type name: audio

   Subtype names: EVRCB1

   Required parameters: none

   Optional parameters:

      ptime: see RFC 4566 [7].

      maxptime: The maximum amount of media that can be encapsulated in
         each packet, expressed as time in milliseconds. The time MUST
         be calculated as the sum of the time the media present in the
         packet represents. The time SHOULD be a multiple of the
         duration of a single codec data frame (20 msec). If not
         signaled, the default maxptime value MUST be 200 milliseconds.

      fixedrate: Indicates the EVRC-B 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 for definition. If this parameter
         is not present, the default value 1 MUST be assumed.

      dtxmax: see Section 6.1

      dtxmin: see Section 6.1

      hangover: see Section 6.1

   Encoding considerations:
      This media type is framed binary data (see RFC 4288, Section 4.8)
      and is defined for transfer of EVRC-B-encoded data via RTP using
      the compact bundled format as described in RFC 4788.

   Security considerations: See Section 9 of RFC 4788.

   Interoperability considerations: none.

   Published specification:
      The EVRC-B vocoder is specified in 3GPP2 C.S0014-B [3]. Transfer
      method with compact bundled RTP format 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:
      Qiaobing Xie <Qiaobing.Xie@motorola.com>

   Intended usage: COMMON

   Restrictions on usage:
      This media type depends on RTP framing; hence, it is only defined
      for transfer via RTP (RFC 3550 [5]). Transfer within other
      framing protocols is not defined at this time.

   Author:
      Qiaobing Xie

   Change controller:
      IETF Audio/Video Transport working group delegated from the IESG.

6.5. Updated Registration of Media Type EVRC

   (The definition is from RFC 3558, added with the optional DTX
   parameters, and updated with the new template specified in [10].)

   Type name: audio

   Subtype names: EVRC

   Required parameters: none

   Optional parameters:

      ptime: Defined as usual for RTP audio (see RFC 4566).

      maxptime: The maximum amount of media that can be encapsulated in
         each packet, expressed as time in milliseconds. The time SHALL
         be calculated as the sum of the time the media present in the
         packet represents. The time SHOULD be a multiple of the
         duration of a single codec data frame (20 msec). If not
         signaled, the default maxptime value SHALL be 200 milliseconds.

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

      silencesupp: see Section 6.1 for definition. If this parameter
         is not present, the default value 1 MUST be assumed.

      dtxmax: see Section 6.1

      dtxmin: see Section 6.1

      hangover: see Section 6.1

   Encoding considerations:
      This media type is framed binary data (see RFC 4288, Section 4.8),
      and is defined for transfer of EVRC-encoded data via RTP using the
      Interleaved/Bundled packet format specified in Sections 4.1, 6,
      and 7 of RFC 3558. It is also defined for other transfer methods
      using the storage format specified in Section 11 of RFC 3558.

   Security considerations: See Section 14, "Security Considerations",
      of RFC 3558.

   Interoperability considerations:
      The DTX parameters are receiver options. Existing RFC 3558
      implementations will not send any of the DTX parameters in their
      SDP and will ignore any DTX parameters they receive. The adaptive
      DTX behavior of DTX-capable EVRC codecs (as detailed in [8],
      Section 4.3.5) ensures interoperability with non-DTX EVRC codecs.

   Published specification:
      The EVRC vocoder is specified in 3GPP2 C.S0014 [2]. Transfer
      methods are specified in RFC 3558.

   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 information applies for storage format only.

         Magic number: #!EVRC\n (see Section 11 of RFC 3558)
         File extensions: evc, EVC
         Macintosh file type code: none
         Object identifier or OID: none

   Person & email address to contact for further information:
      Qiaobing Xie <Qiaobing.Xie@motorola.com>

   Intended usage: COMMON

   Restrictions on usage:
      This media type may be used with RTP framing (RFC 3550 [5]) and as
      a storage format. When used with RTP, the procedures in RFC 3558,
      Section 4.1, MUST be followed. In all other contexts, the storage
      format defined in RFC 3558, Section 11, MUST be used.

   Author:
      Adam Li/Qiaobing Xie

   Change controller:
      IETF Audio/Video Transport working group delegated from the IESG.

6.6. Updated Registration of Media Type EVRC0

   (The definition is from RFC 3558, added with the optional DTX
   parameters, and updated with the new template specified in [10].)

   Type name: audio

   Subtype names: EVRC0

   Required parameters: none

   Optional parameters:

      silencesupp: see Section 6.1 for definition. If this parameter
         is not present, the default value 1 MUST be assumed.

    dtxmax: see Section 6.1

    dtxmin: see Section 6.1

    hangover: see Section 6.1

Encoding considerations:
   This media type is framed binary data (see RFC 4288, Section 4.8)
   and is only defined for transfer of EVRC-encoded data via RTP
   using the Header-Free packet format specified in Section 4.2 of
   RFC 3558.

Security considerations: See Section 14, "Security Considerations",
   of RFC 3558.

Interoperability considerations:
   The DTX parameters are receiver options. Existing RFC 3558
   implementations will not send any of the DTX parameters in their
   SDP and will ignore any DTX parameters they receive. The adaptive
   DTX behavior of DTX-capable EVRC codecs (as detailed in [8],
   Section 4.3.5) ensures interoperability with non-DTX EVRC codecs.

Published specification:
   The EVRC vocoder is specified in 3GPP2 C.S0014 [2]. Transfer
   methods are specified in RFC 3558.

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:
   Qiaobing Xie <Qiaobing.Xie@motorola.com>

Intended usage: COMMON

Restrictions on usage:
   This media type depends on RTP framing; hence, it is only defined
   for transfer via RTP (RFC 3550 [5]). Transfer within other
   framing protocols is not defined at this time.

Author:
   Adam Li/Qiaobing Xie

   Change controller:
      IETF Audio/Video Transport working group delegated from the IESG.

6.7. Mapping MIME Parameters into SDP

   The information carried in the MIME media type specification has a
   specific mapping to fields in the Session Description Protocol (SDP)
   [7], which is commonly used to describe RTP sessions. When SDP is
   used to specify sessions employing the compact bundled format for
   EVRC/EVRC-B-encoded speech, the mapping is as follows:

   o The MIME type ("audio") goes in SDP "m=" as the media name.

   o The MIME subtype ("EVRC", "EVRC0", "EVRC1", "EVRCB", EVRCB0", or
      "EVRCB1") goes in SDP "a=rtpmap" as the encoding name.

   o The optional parameters "ptime" and "maxptime" (for subtypes EVRC,
      EVRC1, EVRCB, and EVRCB1) go in the SDP "a=ptime" and "a=maxptime"
      attributes, respectively.

   o The optional parameter "maxinterleave" (for subtypes EVRC and
      EVRCB) goes in the SDP "a=fmtp" attribute by copying it directly
      from the MIME media type string as "maxinterleave=value".

   o The optional parameter "fixedrate" (for subtypes EVRC1 and EVRCB1)
      goes in the "a=fmtp" attribute by copying it directly from the
      MIME media type string as "fixedrate=value".

   o The optional parameters "silencesupp", "dtxmax", "dtxmin", and
      "hangover" go in the "a=fmtp" attribute by copying it directly
      from the MIME media type string as "silencesupp=value",
      "dtxmax=value", "dtxmin=value", and "hangover=value",
      respectively.

   Example of usage of EVRC1:

     m=audio 49120 RTP/AVP 97
     a=rtpmap:97 EVRC1/8000
     a=fmtp:97 fixedrate=0.5
     a=maxptime:120

   Example of usage of EVRCB:

     m=audio 49120 RTP/AVP 97
     a=rtpmap:97 EVRCB/8000
     a=maxptime:120

   Example of usage of EVRCB0:

     m=audio 49120 RTP/AVP 97
     a=rtpmap:97 EVRCB0/8000

   Example of usage of EVRCB1:

     m=audio 49120 RTP/AVP 97
     a=rtpmap:97 EVRCB1/8000
     a=fmtp:97 fixedrate=0.5
     a=maxptime:100

   Example of usage of EVRC with DTX with silencesupp=1:

     m=audio 49120 RTP/AVP 97
     a=rtpmap:97 EVRC/8000
     a=fmtp:97 silencesupp=1 dtxmax=32 dtxmin=12 hangover=1

   Example of usage of EVRC with DTX with silencesupp=0:

     m=audio 49120 RTP/AVP 97
     a=rtpmap:97 EVRC/8000
     a=fmtp:97 silencesupp=0

6.8. Usage in Offer/Answer

   All SDP parameters in this payload format are declarative, and all
   reasonable values are expected to be supported. In particular, when
   DTX is supported, the RTP sender implementation SHOULD support
   hangover, dtxmin, and dtxmax values from 0 to 255. Thus, the
   standard usage of Offer/Answer, as described in RFC 3264 [6], SHOULD
   be followed.

   In addition, the following rules MUST be followed while negotiating
   DTX parameters:

   1. If any DTX parameter is not present in either offer and/or
       answer, the default value of the DTX parameter MUST be assumed.

   2. If silencesupp is present and set to 0 in either offer or answer,
       the values of all received DTX parameters other than silencesupp
       SHOULD be ignored.

   3. In an offer or answer, the value of dtxmax SHOULD always be
       larger than or equal to the value of dtxmin, regardless of
       whether the values are indicated explicitly or implicitly by
       default. Moreover, if the indicated value of dtxmin is larger

      than that of dtxmax, an RTP sender MUST ignore the indicated
      values and MUST fall back on using the default dtxmin and dtxmax
      values.

7. Backward Compatibility with RFC 3558

   This document adds new optional DTX parameters to the original EVRC
   payload subtypes "EVRC" and "EVRC0" defined in RFC 3558. Since the
   new DTX parameters are receiver options, we expect that the existing
   RFC 3558 implementations will not send any of the DTX parameters in
   their SDP and will ignore any DTX parameters they receive. The
   adaptive DTX behavior of DTX-capable EVRC codecs (as detailed in [8],
   Section 4.3.5) ensures the backward interoperability between the DTX-
   capable EVRC codec and non-DTX EVRC codecs.

8. IANA Considerations

   Four (4) new MIME subtype registrations - "EVRC1", "EVRCB", "EVRCB0",
   and "EVRCB1" - are defined in this document (see Section 6.1 -
   Section 6.4) for EVRC-B and compact bundled payload format support.

   For all the EVRC and EVRC-B RTP payload formats defined in RFC 3558
   [4] and RFC 4788, four additional optional parameters -
   "silencesupp", "dtxmax", "dtxmin", and "hangover" - are defined and
   used in DTX.

   The MIME subtype registrations "EVRC" and "EVRC0", originally defined
   in RFC 3558 [4], are updated with the optional DTX parameters (see
   Sections 6.5 and 6.6).

9. Security Considerations

   Implementations using the payload defined in this specification are
   subject to the security considerations discussed in RFC 3558 [4], RFC
   3550 [5], and any appropriate profile (for example, RFC 3551 [9]).
   This payload does not specify any different security services.

10. Acknowledgements

   The following people have made significant contributions to this
   document (in alphabetical order): Parag Agashe, Jim Ashley,
   Harikishan Desineni, Serafin Diaz, Harinath Garudadri, Gouri
   Johanssen, Ananth Kandhadai, Waqar Mohsin, Ashok Roy, Gino Scribano,
   and Gajinder Singh Vij.

   Special thanks to Colin Perkins, Magnus Westerlund, and Adam Li for
   their careful review and comments that significantly improved the
   quality of this document.

11. References

11.1. Normative References

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

   [2]   "Enhanced Variable Rate Codec, Speech Service Option 3 for
          Wideband Spread Spectrum Digital Systems", 3GPP2 C.S0014,
          January 1997.

   [3]   "Enhanced Variable Rate Codec, Speech Service Option 3 and 68
          for Wideband Spread Spectrum Digital Systems", 3GPP2 C.S0014-B
          v1.0, May 2006.

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

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

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

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

   [8]   "Discontinuous Transmission (DTX) of Speech in cdma2000
          Systems", 3GPP2 C.S0076-0, Version 1.0, December 2005.

11.2. Informative References

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

   [10]   Casner, S., "Media Type Registration of RTP Payload Formats",
          Work in Progress, March 2006.

Authors' Addresses

   Qiaobing Xie
   Motorola, Inc.
   1501 W. Shure Drive, 2-F9
   Arlington Heights, IL 60004
   US

   Phone: +1-847-632-3028
   EMail: Qiaobing.Xie@Motorola.com

   Rohit Kapoor
   Qualcomm Inc.
   US

   Phone: +1-858-845-1161
   EMail: rkapoor@qualcomm.com