Internet Engineering Task Force
INTERNET-DRAFT
draft-ietf-avt-rfc2833bis-02.ps

AVT WG
H. Schulzrinne/S. Petrack
Columbia U./eDial

December 6, 2002
Expires: April 2003

RTP Payload for DTMF Digits, Telephony Tones and Telephony Signals

Status of this Memo

Copyright Notice

Abstract

This memo describes how to carry dual-tone multifrequency (DTMF) signaling, other tone signals and telephony events in RTP packets. This document updates RFC 2833.

1 Introduction

This memo defines two payload formats, one for carrying dual-tone multifrequency (DTMF) digits, other line and trunk signals (Section 3), and a second one for general multi-frequency tones in RTP [1] packets (Section 4). Separate RTP payload formats are desirable since low-rate voice codecs cannot be guaranteed to reproduce these tone signals accurately enough for automatic recognition. Defining separate payload formats also permits higher redundancy while maintaining a low bit rate.

The payload formats described here may be useful in at least three applications: DTMF handling for gateways and end systems, as well as "RTP trunks". In the first application, the Internet telephony gateway detects DTMF on the incoming circuits and sends the RTP payload described here instead of regular audio packets. The gateway likely has the necessary digital signal processors and algorithms, as it often needs to detect DTMF, e.g., for two-stage dialing. Having the gateway detect tones relieves the receiving Internet end system from having to do this work and also avoids that low bit-rate codecs like G.723.1 render DTMF tones unintelligible. Secondly, an Internet end system such as an "Internet phone" can emulate DTMF functionality without concerning itself with generating precise tone pairs and without imposing the burden of tone recognition on the receiver.

In the "RTP trunk" application, RTP is used to replace a normal circuit-switched trunk between two nodes. This is particularly of interest in a telephone network that is still mostly circuit-switched. In this case, each end of the RTP trunk encodes audio channels into the appropriate encoding, such as G.723.1 or G.729. However, this encoding process destroys in-band signaling information which is carried using the least-significant bit ("robbed bit signaling") and may also interfere with in-band signaling tones, such as the MF digit tones. In addition, tone properties such as the phase reversals in the ANSam tone, will not survive speech coding. Thus, the gateway needs to remove the in-band signaling information from the bit stream. It can now either carry it out-of-band in a signaling transport mechanism yet to be defined, or it can use the mechanism described in this memorandum. (If the two trunk end points are within reach of the same media gateway controller, the media gateway controller can also handle the signaling.) Carrying it in-band may simplify the time synchronization between audio packets and the tone or signal information. This is particularly relevant where duration and timing matter, as in the carriage of DTMF signals.

1.1 Terminology

In this document, the key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" are to be interpreted as described in RFC 2119 [2] and indicate requirement levels for compliant implementations.

2 Events vs. Tones

A gateway has two options for handling DTMF digits and events. First, it can simply measure the frequency components of the voice band signals and transmit this information to the RTP receiver (Section 4). In this mode, the gateway makes no attempt to discern the meaning of the tones, but simply distinguishes tones from speech signals.

All tone signals in use in the PSTN and meant for human consumption are sequences of simple combinations of sine waves, either added or modulated. (There is at least one tone, the ANSam tone [3] used for indicating data transmission over voice lines, that makes use of periodic phase reversals.)

As a second option, a gateway can recognize the tones and translate them into a name, such as ringing or busy tone. The receiver then produces a tone signal or other indication appropriate to the signal. Generally, since the recognition of signals often depends on their on/off pattern or the sequence of several tones, this recognition can take several seconds. On the other hand, the gateway may have access to the actual signaling information that generates the tones and thus can generate the RTP packet immediately, without the detour through acoustic signals.

In the phone network, tones are generated at different places, depending on the switching technology and the nature of the tone. This determines, for example, whether a person making a call to a foreign country hears her local tones she is familiar with or the tones as used in the country called.

For analog lines, *dial tone* is always generated by the local switch. ISDN terminals may generate dial tone locally and then send a Q.931 SETUP message containing the dialed digits. If the terminal just sends a SETUP message without any Called Party digits, then the switch does digit collection, provided by the terminal as KEYPAD messages, and provides dial tone over the B-channel. The terminal can either use the audio signal on the B-channel or can use the Q.931 messages to trigger locally generated dial tone.

Ringing tone (also called ringback tone) is generated by the local switch at the callee, with a one-way voice path opened up as soon as the callee's phone rings. (This reduces the chance of clipping the called party's response just after answer. It also permits pre-answer announcements or in-band call-progress indications to reach the caller before or in lieu of a ringing tone.) *Congestion tone* and *special information tones* can be generated by any of the switches along the way, and may be generated by the caller's switch based on ISUP messages received. *Busy tone* is generated by the caller's switch, triggered by the appropriate ISUP message, for analog instruments, or the ISDN terminal.

Gateways which send signalling events via RTP MAY send both named signals (Section 3) and the tone representation (Section 4) as a single RTP session, using the redundancy mechanism defined in Section 3.7

to interleave the two representations. It is generally a good idea to send both, since it allows the receiver to choose the appropriate rendering.

If a gateway cannot present a tone representation, it SHOULD also send the audio tones as regular RTP audio packets using either the codec used for regular speech signals or a codec that is known to carry such signals successfully (e.g., PCMU).

> Some low-rate codecs cannot accurately represent certain tones, such as DTMF.

3 RTP Payload Format for Named Telephone Events

3.1 Introduction

The payload format for named telephone events described below is suitable for both gateway and end-to-end scenarios. In the gateway scenario, an Internet telephony gateway connecting a packet voice network to the PSTN recreates the DTMF tones or other telephony events and injects them into the PSTN. Since, for example, DTMF digit recognition takes several tens of milliseconds, the first few milliseconds of a digit will arrive as regular audio packets. Thus, careful time and power (volume) alignment between the audio samples and the events is needed to avoid generating spurious digits at the receiver.

DTMF digits and named telephone events are carried as part of the audio stream, and MUST use the same sequence number and time-stamp base as the regular audio channel to simplify the generation of audio waveforms at a gateway. The default clock frequency is 8,000 Hz, but the clock frequency can be redefined when assigning the dynamic payload type. Named telephone events can be considered a very highly-compressed audio codec, and is treated the same as other codecs.

The payload format described here achieves a higher redundancy even in the case of sustained packet loss than the method proposed for the *Voice over Frame Relay Implementation Agreement* [25].

If an end system is directly connected to the Internet and does not need to generate tone signals again, time alignment and power levels are not relevant. These systems rely on PSTN gateways or Internet end systems to generate DTMF events and do not perform their own audio waveform analysis. An example of such a system is an Internet interactive voice-response (IVR) system.

In circumstances where exact timing alignment between the audio stream and the DTMF digits or other events is not important and data is sent unicast, such as the IVR example mentioned earlier, it may be preferable to use a reliable control protocol rather than RTP packets. In those circumstances, this payload format would not be used.

3.2 Simultaneous Generation of Audio and Events

A source can choose between four approaches:

Events and audio: The sends events and encoded audio packets (e.g., PCMU or the codec used for speech signals) for the same time instant. In that mode, events are treated as redundant encodings for the encoded audio stream.

Events only: The source does not send encoded audio while event tones are active and only sends named events, without any redundancy beyond the periodic updates of longer-lasting events.

Events only, with redundancy: The source does not send encoded audio while event tones are active. It only sends named events, but uses RFC 2198 [4] redundancy, with named events as both primary and redundant encodings.

Events and audio, with redundancy: During an event, the source sends both named events and audio, using RFC 2198 to interleave audio data, current and redundant named events.

Note that a period covered by a named event may overlap in time with a period of audio encoded by other means. This is likely to occur at the onset of a tone and is necessary to avoid possible errors in the interpretation of the reproduced tone at the remote end. Implementations supporting this payload format must be prepared to handle the overlap. It is RECOMMENDED that gateways only render the encoded tone since the audio may contain spurious tones introduced by the audio compression algorithm. However, it is anticipated that these extra tones in general should not interfere with recognition at the far end.

3.3 Event Types

This payload format is used for five different types of signals:

- DTMF tones (Section 3.10);
- fax-related tones (Section 3.11);
- standard subscriber line tones (Section 3.12);
- country-specific subscriber line tones (Section 3.13) and;
- trunk events (Section 3.14).

A compliant implementation MUST support the events listed in Table 1 with the exception of "flash". If it uses some other, out-of-band mechanism for signaling line conditions, it does not have to implement events other than those in Table 1.

In some cases, an implementation may simply ignore certain events, such as fax tones, that do not make sense in a particular environment. Section 3.9 specifies how an implementation can use the SDP "fmtp" parameter within an SDP description to indicate its inability to understand a particular event or range of events.

Depending on the available user interfaces, an implementation MAY render all tones in Table 5 the same or, preferably, use the tones conveyed by the concurrent "tone" payload or other RTP audio payload. Alternatively, it MAY provide a textual representation.

Note that end systems that emulate telephones only need to support the events described in Sections 3.10 and 3.12, while systems that receive trunk signaling need to implement those in Sections 3.10, 3.11, 3.12 and 3.14, since MF trunks also carry most of the "line" signals. Systems that do not support fax or modem functionality do not need to render fax-related events described in Section 3.11.

The RTP payload format is designated as "telephone-event", the MIME type as "audio/telephone-event". The default timestamp rate is 8000 Hz, but other rates may be defined. In accordance with current practice, this payload format does not have a static payload type number, but uses a RTP payload type number established dynamically and out-of-band.

3.4 Use of RTP Header Fields

Timestamp: The RTP timestamp reflects the measurement point for the current packet. The event duration described in Section 3.5 extends forwards from that time. For events that span multiple RTP packets, the RTP timestamp identifies the beginning of the event, i.e., several RTP packets may carry the same timestamp. For long-lasting events that have to be split into subevents (see below), the timestamp indicates the beginning of the subevent. If there are multiple events in one RTP packet, the events MUST be contiguous.

The receiver calculates jitter for RTCP receiver reports based on all packets with a given timestamp. Note: The jitter value should primarily be used as a means for comparing the reception quality between two users or two time-periods, not as an absolute measure.

Marker bit: The RTP marker bit indicates the beginning of a new event. If an event lasts more than the maximum time representable by the duration field (see below), the event is split into subevents. Only the first one will have the marker bit set.

3.5 Payload Format

The payload format is shown in Fig. 1.

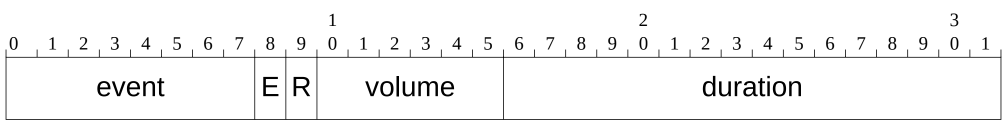

Figure 1: Payload Format for Named Events

events: The events are encoded as shown in Sections 3.10 through 3.14.

volume: For DTMF digits and other events representable as tones, this field describes the power level of the tone, expressed in dBm0 after dropping the sign. Power levels range from 0 to -63 dBm0. Thus, larger values denote lower volume. This value is defined only for events containing the value "yes" in the "volume?" column of tables 1 to 5 and MUST be set to zero for other events. If a zero volume is indicated for an event for which the volume field is defined, then the receiver MAY reconstruct the volume from the volume of non-event audio or MAY use the nominal value specified by the ITU Recommendation or other document defining the tone.

This is backwards compatibility with RFC 2833, where the volume field was defined only for DTMF events.

duration: Duration of this event, in timestamp units, expressed as an unsigned integer. For a non-zero value, the event began at the instant identified by the RTP timestamp and has so far lasted as long as indicated by this parameter. The event may or may not have ended. If the event duration exceeds the maximum representable by the duration field, the event is split into several contiguous subevents, where all but the last event have the maximum duration expressible in the duration field (0xFFFF). The receiver uses the absence of a gap between the events to detect that it is receiving a single long-duration event. Only the first subevent will begin with a marker bit and only the last subevent will end with the "E" bit set (see below).

The special duration value of zero is reserved to indicate that the event lasts “forever”, i.e., is a state and is considered to be effective until updated. Only events marked with a number in the “state?” column the tables below are allowed to use a duration value of zero. Other named events with such a duration SHOULD be ignored. The state number indicates which states are mutually exclusive. Among all states with the same state number, only one can be active. (For example, the “on hook” state automatically clears the “off hook” state and any of the ABCD states clear the previous ABCD state.)

Events marked as states MAY use a non-zero duration, indicating that the sender intends to refresh the state before the time duration has elapsed (“soft state”). For robustness, the sender SHOULD retransmit “state” event periodically.

For a sampling rate of 8000 Hz, this field is sufficient to express event durations of up to approximately 8 seconds.

E: If set to a value of one, the “end” bit indicates that this packet contains the end of the event. Thus, the duration parameter, if non-zero, measures the complete duration of the event unless the event was concatenated from multiple very long subevents where all but the last had a duration value of 0xFFFF.

A sender MAY delay setting the end bit until retransmitting the last packet for a tone, rather than on its first transmission. This avoids having to wait to detect whether the tone has indeed ended.

Some events are actually states, i.e., the appearence of a different named event implies the end of the previous state. Thus, for named RTP events labeled “state” in Tables 1 through 5, sending of a packet with the “E” bit set is OPTIONAL. State events are sent with zero duration.

Receiver implementations MAY use different algorithms to create tones, including the two described here. In the first, the receiver simply places a tone of the given duration in the audio playout buffer at the location indicated by the timestamp. As additional packets are received that extend the same tone, the waveform in the playout buffer is extended accordingly. (Care has to be taken if audio is mixed, i.e., summed, in the playout buffer rather than simply copied.) Thus, if a packet in a tone lasting longer than the packet interarrival time gets lost and the playout delay is short, a gap in the tone may occur.

Alternatively, the receiver can start a tone and play it until it receives a packet with the “E” bit set, the next tone, distinguished by a different timestamp value or a given time period elapses. This is more robust against packet loss, but may extend the tone beyond its original duration if all retransmissions of the last packet in an event are lost. Limiting the time period of extending the tone is necessary to avoid that a tone “gets stuck”. This algorithm is not a license for senders to set the duration field to zero; it MUST be set to the current duration as described, since this is needed to create accurate events if the first event packet is lost, among other reasons.

Regardless of the algorithm used, the tone SHOULD NOT be extended by more than three packet interarrival times. A slight extension of tone durations and shortening of pauses is generally harmless.

R: This field is reserved for future use. The sender MUST set it to zero, the receiver MUST ignore it.

3.6 Sending Event Packets

An audio source SHOULD start transmitting event packets as soon as it recognizes an event. A source has wide latitude as to how often it sends event updates afterwards. A natural interval is the spacing between audio packets. (Recall that a single RTP packet can contain multiple audio frames for frame-based codecs and that the packet interval can vary during a session.) Alternatively, a source MAY decide to use a different spacing for event updates, with a value of 50 ms RECOMMENDED. A receiver should not rely on a particular event packet spacing. Timing information is contained in the RTP timestamp, allowing precise recovery of inter-event times. Thus, the sender does not need to maintain precise or consistent time intervals between event packets. in order to maintain precise inter-event times,

Q.24 [5], Table A-1, indicates that all administrations surveyed use a minimum signal duration of 40 ms, with signaling velocity (tone and pause) of no less than 93 ms.

If an event continues for more than one period, the source generating the events should send a new event packet with the RTP timestamp value corresponding to the beginning of the event and the duration of the event increased by the elapsed period. If an event has ended and there has been no new event in the last interval, the event SHOULD be retransmitted three times at the interval used by the source for updates. (If a new event is recognized during the retransmissions and RFC 2198 is in use, the old event will be part of the redundancy.) This ensures that the duration of the event can be recognized correctly even if the last packet for an event is lost. The last update is also repeated at the end of subevents, as there is no other way to detect the duration of the subevent.

In all cases, the RTP sequence number MUST be incremented by one in each RTP packet.

DTMF digits and events are sent incrementally to avoid having the receiver wait for the completion of the event. Since some tones are two seconds long, this would incur a substantial delay. The transmitter does not know if event length is important and thus needs to transmit immediately and incrementally. If the receiver application does not care about event length, the incremental transmission mechanism avoids delay. Some applications, such as gateways into the PSTN, care about both delays and event duration.

For events with a duration shorter than a typical packet interval, for example, V.21 bits (Section 3.11), it is RECOMMENDED that multiple events are represented by a single RFC 2198 [4] packet, as described in Section 3.7.

Multiple named events can be packet into a single RTP packet if and only if the events are contiguous, i.e., occur without pause, and if the last event packed into a packet occurs fast enough to avoid excessive delays at the receiver.

This approach is similar to having multiple frames of frame-based audio in one RTP packet.

3.7 Reliability

During an event, the RTP event payload format provides incremental updates on the event. The error resiliency depends on the playout delay at the receiver. For example, for a playout delay of 120 ms and a packet gap of 50 ms, two packets in a row can get lost without causing a gap in the tones generated at the receiver.

The audio redundancy mechanism described in RFC 2198 [4] MAY be used to recover from packet loss across events. The effective data rate is r times 64 bits (32 bits for the redundancy header and 32 bits for the telephone-event payload) every 50 ms or r times 1280 bits/second, where r is the number of redundant events carried in each packet. The value of r is an implementation trade-off, with a value of 5 suggested.

The timestamp offset in this redundancy scheme has 14 bits, so that it allows a single packet to "cover" 2.048 seconds of telephone events at a sampling rate of 8000 Hz. Including the starting time of previous events allows precise reconstruction of the tone sequence at a gateway. The scheme is resilient to consecutive packet losses spanning this interval of 2.048 seconds or r digits, whichever is less. Note that for previous digits, only an average loudness can be represented.

An encoder MAY treat the event payload as a highly-compressed version of the current audio frame. In that mode, each RTP packet during an event would contain the current audio codec rendition (say, G.723.1

or G.729) of this digit as well as the representation described in Section 3.5, plus any previous events seen earlier.

> This approach allows dumb gateways that do not understand this format to function. See also the discussion in Section 1.

3.8 Example

A typical RTP packet, where the user is just dialing the last digit of the DTMF sequence “911”. The first digit was 200 ms long (1600 timestamp units) and started at time 0, the second digit lasted 250 ms (2000 timestamp units) and started at time 800 ms (6400 timestamp units), the third digit was pressed at time 1.4 s (11,200 timestamp units) and the packet shown was sent at 1.45 s (11,600 timestamp units). The frame duration is 50 ms. To make the parts recognizable, the figure below ignores byte alignment. Timestamp and sequence number are assumed to have been zero at the beginning of the first digit. In this example, the dynamic payload types 96 and 97 have been assigned for the redundancy mechanism and the telephone event payload, respectively.

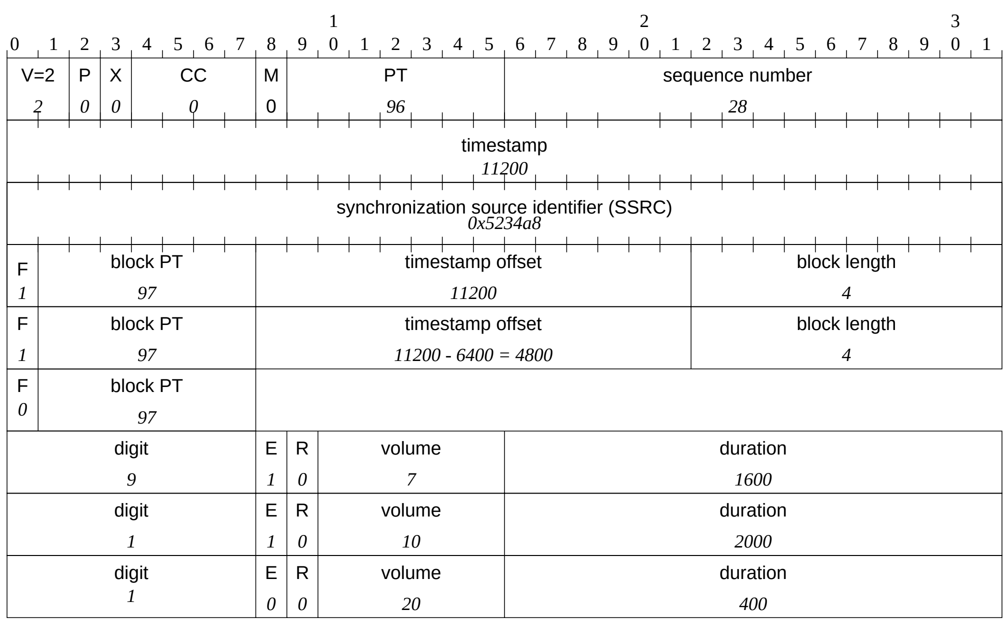

Figure 2: Example RTP packet after dialing “911”

3.9 Indication of Receiver Capabilities using SDP

Receivers MAY indicate which named events they can handle, for example, by using the Session Description Protocol (RFC 2327 [6]). SDP descriptions using the event payload MUST contain a fmtp format attribute that lists the event values that the receiver can process:

```
a=fmtp:<format> <list of values>
```

Event	encoding (decimal)	state?	volume?
0–9	0–9		yes
*	10		yes
#	11		yes
A–D	12–15		yes
Flash	16		no

Table 1: DTMF named events

The list of values consists of comma-separated elements, which can be either a single decimal number or two decimal numbers separated by a hyphen (dash), where the second number is larger than the first. No whitespace is allowed between numbers or hyphens. The list does not have to be sorted.

For example, if the payload format uses the payload type number 100, and the implementation can handle the DTMF tones (events 0 through 15) and the dial and ringing tones, it would include the following description in its SDP message:

```
a=fmtp:100 0-15,66,70
```

The corresponding MIME parameter is “events”, so that the following sample media type definition corresponds to the SDP example above:

```
audio/telephone-event;events="0-15,66,67";rate="8000"
```

3.10 DTMF Events

Tables 1 summarizes the DTMF-related named events within the telephone-event payload format. The “volume?” column indicates whether the receiver should interpret the volume indication.

The “Flash” event must only be sent when the state is “Off Hook”.

3.11 Data Modem and Fax Events

Table 3.11 summarizes the events and tones that can appear on a subscriber line serving a fax machine or modem. The tones are described below, with additional detail in Table 7.

ANS: This 2100 +/- 15 Hz tone is used to disable echo suppression for data transmission [7, 8]. For fax machines, Recommendation T.30 [8] refers to this tone as called terminal identification (CED) answer tone.

/ANS: This is the same signal as ANS, except that it reverses phase at an interval of 450 +/- 25 ms. It disables both echo cancellers and echo suppressors. (In the ITU Recommendation V.25 [7, Fig. 3], an ANS with a bar on top refers to individual phase-reversed cycles rather than to the entire signal.)

ANSam: The modified answer tone (ANSam) [3, Section 7.2] is a sinewave signal at 2100 +/- 1 Hz *without* phase reversals, amplitude-modulated by a sinewave at 15 +/- 0.1 Hz. This tone is sent by modems if network echo canceller disabling is not required.

/ANSam: The modified answer tone with phase reversals (ANSam) [3] is a sinewave signal at 2100 +/- 1 Hz with phase reversals at intervals of 450 +/- 25 ms, amplitude-modulated by a sinewave at 15 +/- 0.1 Hz. This tone [9, 7] is sent by modems [10] and faxes to disable echo suppressors.

These definitions of the ANS, /ANS, ANSam and /ANSam tones refer to the entire signal. Unlike ITU Recommendation V.25 [7], they do not refer to individual 450 ms cycles.

An ANS or ANSam event packet should not be sent until it is possible to discriminate between an ANS and ANSam event. It is however, permissible to send an ANS or ANSam event packet before phase reversals can be detected. Phase reversals, if any, occur at intervals of 450 +/- 25 ms. If a phase reversal is detected after an ANS or ANSam event packet is sent, it must be followed by the transmission of an /ANS or /ANSam event packet.

CNG: After dialing the called fax machine's telephone number (and before it answers), the calling Group III fax machine (optionally) begins sending a CalliNG tone (CNG) consisting of an interrupted tone of 1100 Hz. [8]

CRdi: Capabilities Request (CRd), initiating side, [11] is a dual-tone signal with tones at 1375 Hz and 2002 Hz for 400 ms, followed by a single tone at 1900 Hz for 100 ms. "This signal requests the remote station transition from telephony mode to an information transfer mode and requests the transmission of a capabilities list message by the remote station. In particular, CRdi is sent by the initiating station during the course of a call, or by the calling station at call establishment in response to a CRe or MRe."

CRdr: CRdr is the response tone to CRdi (see above). It consists of a dual-tone signal with tones at 1529 Hz and 2225 Hz for 400 ms, followed by a single tone at 1900 Hz for 100 ms.

CRe: Capabilities Request (CRe) [11] is a dual-tone signal with tones at tones at 1375 Hz and 2002 Hz for 400 ms, followed by a single tone at 400 Hz for 100 ms. "This signal requests the remote station transition from telephony mode to an information transfer mode and requests the transmission of a capabilities list message by the remote station. In particular, CRe is sent by an automatic answering station at call establishment."

CT: "The calling tone [7] consists of a series of interrupted bursts of binary 1 signal or 1300 Hz, on for a duration of not less than 0.5 s and not more than 0.7 s and off for a duration of not less than 1.5 s and not more than 2.0 s." Modems not starting with the V.8 call initiation tone often use this tone.

ESi: Escape Signal (ESi) [11] is a dual-tone signal with tones at 1375 Hz and 2002 Hz for 400 ms, followed by a single tone at 980 Hz for 100 ms. "This signal requests the remote station transition from telephony mode to an information transfer mode. signal ESi is sent by the initiating station."

ESr: Escape Signal (ESr) [11] is a dual-tone signal with tones at 1529 Hz and 2225 Hz for 400 ms, followed by a single tone at 1650 Hz for 100 ms. Same as ESi, but sent by the responding station.

MRdi: Mode Request (MRd), initiating side, [11] is a dual-tone signal with tones at 1375 Hz and 2002 Hz for 400 ms followed by a single tone at 1150 Hz for 100 ms. "This signal requests the remote station transition from telephony mode to an information transfer mode and requests the transmission of a mode select message by the remote station. In particular, signal MRd is sent by the initiating station during the course of a call, or by the calling station at call establishment in response to an MRe." [11]

Procedure	indications
V.25 and V.8	ANS
V.25, echo canceller disabled	ANS, /ANS, ANS, /ANS
V.8	ANSam
V.8, echo canceller disabled	/ANSam

Table 2: Use of ANS, ANSam and /ANSam in V.x recommendations

MRdr: MRdr is the response tone to MRdi (see above). It consists of a dual-tone signal with tones at 1529 Hz and 2225 Hz for 400 ms, followed by a single tone at 1150 Hz for 100 ms.

MRe: Mode Request (MRe) [11] is a dual-tone signal with tones at 1375 Hz and 2002 Hz for 400 ms, followed by a single tone at 650 Hz for 100 ms. "This signal requests the remote station transition from telephony mode to an information transfer mode and requests the transmission of a mode select message by the remote station. In particular, signal MRe is sent by an automatic answering station at call establishment." [11]

V.21: V.21 describes a 300 b/s full-duplex modem that employs frequency shift keying (FSK). It is used by Group 3 fax machines to exchange T.30 information. The calling transmits on channel 1 and receives on channel 2; the answering modem transmits on channel 2 and receives on channel 1. Each bit value has a distinct tone, so that V.21 signaling comprises a total of four distinct tones.

ANS2225: This 2225 Hz answer tone is described in ITU Recommendation V.18, Annex D [12] for one of several classes of modems operating in the text telephone mode. It is also referred to in ITU Recommendation V.22 [13]. This is a pure tone with no amplitude modulation and no semantics attached to phase reversals, if there are any.

> Initially a proprietary "Bell System" method, the 2225 Hz answer tone is now included in ITU V.18, Annex D which addresses TDD (telecommunications for the disabled) equipment. It is necessary to accommodate it for completeness, and for compliance with various legal ordinances. A distinct number must be allocated to this event since it must be differentiated from the normal, 2100 Hz answer tone when reproduced at the far-end gateway.

CI: CI (call indicator) [3]. It is also used by V.18 [12]. It consists of 10 V.21 "1" bits. To fully express the call indicator, it would be followed by a call function octet, composed of individual V.21 bit events.

V.21 flag: The V.21 preamble flag consists of one second of HDLC flag octets (0x7E). Fax machines send it after detecting a T.30 preamble. (Note that devices can start sending the event as soon as they detect the tone; they do not have to wait until the end of the flag event.)

In summary, procedures in Table 2 are used.

3.12 Line Events

Table 4 summarizes the events and tones that can appear on a subscriber line.

ITU Recommendation E.182 [14] defines when certain tones should be used. It defines the following standard tones that are heard by the caller:

Event	encoding (decimal)	state?	volume?
Answer tone (ANS)	32		yes
/ANS	33		yes
ANSam	34		yes
/ANSam	35		yes
Calling tone (CNG)	36		yes
V.21 channel 1, “0” bit	37		yes
V.21 channel 1, “1” bit	38		yes
V.21 channel 2, “0” bit	39		yes
V.21 channel 2, “1” bit	40		yes
CRdi	41		yes
CRdr	42		yes
CRe	43		yes
ESi	44		yes
ESr	45		yes
MRdi	46		yes
MRdr	47		yes
MRe	48		yes
CT	49		yes
ANS2225	52		yes
CI	53		yes
V.21 preamble flag	54		yes

Table 3: Data and fax named events

Dial tone: The exchange is ready to receive address information.

PABX internal dial tone: The PABX is ready to receive address information.

Special dial tone: Same as dial tone, but the caller's line is subject to a specific condition, such as call diversion or a voice mail is available (e.g., "stutter dial tone").

Second dial tone: The network has accepted the address information, but additional information is required.

Ring: This named signal event causes the recipient to generate an alerting signal ("ring"). The actual tone or other indication used to render this named event is left up to the receiver. (This differs from the ringing tone, below, heard by the *caller*.)

Ringing tone: The call has been placed to the callee and a calling signal (ringing) is being transmitted to the callee. This tone is also called "ringback".

Special ringing tone: A special service, such as call forwarding or call waiting, is active at the called number.

Busy tone: The called telephone number is busy.

Congestion tone: Facilities necessary for the call are temporarily unavailable.

Calling card service tone: The calling card service tone consists of 60 ms of the sum of 941 Hz and 1477 Hz tones (DTMF '#'), followed by 940 ms of 350 Hz and 440 Hz (U.S. dial tone), decaying exponentially with a time constant of 200 ms.

Special information tone: The callee cannot be reached, but the reason is neither "busy" nor "congestion". This tone should be used before all call failure announcements, for the benefit of automatic equipment.

Comfort tone: The call is being processed. This tone may be used during long post-dial delays, e.g., in international connections.

Hold tone: The caller has been placed on hold.

Record tone: The caller has been connected to an automatic answering device and is requested to begin speaking.

Caller waiting tone: The called station is busy, but has call waiting service.

Pay tone: The caller, at a payphone, is reminded to deposit additional coins.

Positive indication tone: The supplementary service has been activated.

Negative indication tone: The supplementary service could not be activated.

Off-hook warning tone: The caller has left the instrument off-hook for an extended period of time.

The following tones can be heard by either calling or called party during a conversation:

Call waiting tone: Another party wants to reach the subscriber.

Event	encoding (decimal)	state?	volume?
Off Hook	64	64	no
On Hook	65	64	no
Dial tone	66		yes
PABX internal dial tone	67		yes
Special dial tone	68		yes
Second dial tone	69		yes
Ringing tone	70		yes
Special ringing tone	71		yes
Busy tone	72		yes
Congestion tone	73		yes
Special information tone	74		yes
Comfort tone	75		yes
Hold tone	76		yes
Record tone	77		yes
Caller waiting tone	78		yes
Call waiting tone	79		yes
Pay tone	80		yes
Positive indication tone	81		yes
Negative indication tone	82		yes
Warning tone	83		yes
Intrusion tone	84		yes
Calling card service tone	85		yes
Payphone recognition tone	86		yes
CPE alerting signal (CAS)	87		yes
Off-hook warning tone	88		yes
Ring	89		yes

Table 4: E.182 line events

Warning tone: The call is being recorded. This tone is not required in all jurisdictions.

Intrusion tone: The call is being monitored, e.g., by an operator.

CPE alerting signal (CAS): A tone used to alert a device to an arriving in-band FSK data transmission. A CPE alerting signal is a combined 2130 and 2750 Hz tone, both with tolerances of 0.5% and a duration of 80 to 85 ms. The CPE alerting signal is used with ADSI services and Call Waiting ID services [15].

The following tones are heard by operators:

Payphone recognition tone: The person making the call or being called is using a payphone (and thus it is ill-advised to allow collect calls to such a person).

3.13 Extended Line Events

Table 5 summarizes country-specific events and tones that can appear on a subscriber line.

Event	encoding (decimal)	state?
Acceptance tone	96	
Confirmation tone	97	
Dial tone, recall	98	
End of three party service tone	99	
Facilities tone	100	
Line lockout tone	101	
Number unobtainable tone	102	
Offering tone	103	
Permanent signal tone	104	
Preemption tone	105	
Queue tone	106	
Refusal tone	107	
Route tone	108	
Valid tone	109	
Waiting tone	110	
Warning tone (end of period)	111	
Warning Tone (PIP tone)	112	

Table 5: Country-specific Line events

3.14 Trunk Events

Table 6 summarizes the events and tones that can appear on a trunk. Trunks can also carry line events (Section 3.12), since multi-frequency (MF) signaling does not include backward signals [26] (p. 93) used outside the United States in MFC signaling systems such as MFC-R2 [16]. Unfortunately, frequency pairs with the frequency 1,700 Hz have many different names, depending on which signaling system they are used for. The terms Code 11, Code 12, KP1, KP2, and ST are found in Q.151 [17] describing Signaling System 5 (SS5).

The frequencies for the (United States) MF signals are as follows:

Hz	900	1100	1300	1500	1700
700	1	2	4	7	Code 11, ST3P/KP3P
900		3	5	8	Code 12, ST1P/KP1P
1100			6	9	KP1, KP
1300				0	KP2, S2, ST2, ST2P, KP2P
1500					ST, S0

ITU-T R2 MFC tones [16] are composed of the following frequencies (in Hz):

Signal number	Forward Backward	1380 1140	1500 1020	1620 900	1740 780	1860 660	1980 540
1		X	X				
2		X		X			
3			X	X			
4		X			X		
5			X		X		
6				X	X		
7		X				X	
8			X			X	
9				X		X	
10					X	X	
11		X					X
12			X				X
13				X			X
14					X		X
15						X	X

Event	encoding (decimal)	state?	volume?
MF 0…9	128…137		yes
MF K0 or KP	138		yes
MF R1 KP1 (1100/1700 Hz)	139		yes
MF R1 KP2 (1300/1700 Hz)	140		yes
MF R1 ST (end-of-pulsing) (1500/1700 Hz)	141		yes
MF R1 S1 (900/1700 Hz)	142		yes
Reserved	143		yes
ABCD signaling (see below)	144…159	144	no
Reserved	160…166		
Continuity tone (2010 Hz)	167		yes
Continuity tone (1780 Hz)	168		yes
Reserved	169…173		
MF S3	174		yes
Trunk unavailable	175		no
MFC Forward 1…15	176…190		yes
MFC Backward 1…15	191…205		yes

Table 6: Trunk events

ABCD transitional: 4-bit signaling used by digital trunks. For N-state ($N < 16$) signaling, the first N values are used. ABCD signaling events are all mutually exclusive states. The most recent state transition determines the current state.

The T1 ESF (extended super frame format) allows 2, 4, and 16 state signalling bit options. These signalling bits are named A, B, C, and D. Signalling information is sent as robbed bits in frames 6, 12, 18, and 24 when using ESF T1 framing. A D4 superframe only transmits 4-state signalling with A and B bits. On the CEPT E1 frame, all signalling is carried in timeslot 16, and two channels of 16-state (ABCD) signalling are sent per frame.

Since this information is a state rather than a changing signal, implementations SHOULD use the following triple-redundancy mechanism, similar to the one specified in ITU-T Rec. I.366.2 [18], Annex L. At the time of a transition, the same ABCD information is sent 3 times at an interval of 5 ms. If another transition occurs during this time, then this continues. After a period of no change, the ABCD information is sent every 5 seconds.

Continuity tones: Tones used for testing circuit continuity. A tone of 1780 Hz is sent by the calling exchange. If received by the called exchange, it returns a “continuity verified” tone of 2010 Hz.

MFC R2 signaling: R2 signaling is a compound of line, continuous, out-of-band, link by link, channel associated signaling and (inter)register, multifrequency, compelled, in-band, end to end, channel associated signaling. Line part of R2 signaling, [19], may be analog (or one-bit, A bit in 16th channel, [27]) version (R2A, [20]) and/or digital (two-bit, A and B bits) version (R2D, [21]). In R2 signaling, the signaling sequence is initiated from the outgoing exchange by sending a line “seizing” signal. After line “seizing” signal (and “seizing acknowledgment” signal in R2D) signaling sequence continues by MF signals. Forward MF signals belong to Groups I and II [16]. Backward MF signals belong to Groups A and B [16].

R2 is a compelled tone signaling protocol, meaning that one tone is played until an “acknowledgment or directive for the next tone” is received which indicates that the original tone should cease. In R2 signaling, the signaling sequence is initiated from the outgoing exchange by sending a forward Group I signal. The first forward signal is typically the first digit of the called number. The incoming exchange typically replies with a backward Group A-1 indicating to the outgoing exchange to send the next digit of the called number.

The tones have meaning, however, the meaning varies depending on where the tone occurs in the signaling. The meaning may also depend on the country. Thus, to avoid an unmanageable number of events, this document simply provides means to indicate the 15 forward and 15 backward MF R2 tones.

Trunk unavailable: The trunk is unavailable for service. The length of the downtime is indicated in the duration field. The duration field is set to a value that allows adequate granularity in describing downtime. A value of 1 second is RECOMMENDED. When the trunk becomes unavailable, this event is sent with the same timestamp three times at an interval of 20 ms. If the trunk persists in the unavailable state at the end of the indicated duration, then it is retransmitted, preferably with the same redundancy scheme.

Unavailability of the trunk might result from a failure or an administrative action. This event is used in a stateless manner to synchronize trunk unavailability between equipment connected through provisioned RTP trunks. It avoids the unnecessary consumption of bandwidth in sending a continuous stream of RTP packets with a fixed payload for the duration of the downtime, as would be required in certain E1-based applications. In T1-based applications, trunk conditioning via the ABCD transitional events can be used instead.

4 RTP Payload Format for Telephony Tones

4.1 Introduction

As an alternative to describing tones and events by name, as described in Section 3, it is sometimes preferable to describe them by their waveform properties. In particular, recognition is faster than for naming signals since it does not depend on recognizing durations or pauses.

There is no single international standard for telephone tones such as dial tone, ringing (ringback), busy, congestion ("fast-busy"), special announcement tones or some of the other special tones, such as payphone recognition, call waiting or record tone. However, across all countries, these tones share a number of characteristics [22]:

- Telephony tones consist of either a single tone, the addition of two or three tones or the modulation of two tones. (Almost all tones use two frequencies; only the Hungarian "special dial tone" has three.) Tones that are mixed have the same amplitude and do not decay.
- Tones for telephony events are in the range of 25 (ringing tone in Angola) to 1800 Hz. CED is the highest used tone at 2100 Hz. The telephone frequency range is limited to 3,400 Hz. (The piano has a range from 27.5 to 4186 Hz.)
- Modulation frequencies range between 15 (ANSam tone) to 480 Hz (Jamaica). Non-integer frequencies are used only for frequencies of 16 2/3 and 33 1/3 Hz. (These fractional frequencies appear to be derived from older AC power grid frequencies.)
- Tones that are not continuous have durations of less than four seconds.
- ITU Recommendation E.180 [23] notes that different telephone companies require a tone accuracy of between 0.5 and 1.5%. The Recommendation suggests a frequency tolerance of 1%.

4.2 Examples of Common Telephone Tone Signals

As an aid to the implementor, Table 7 summarizes some common tones. The rows labeled "ITU . . . " refer to the general recommendation of Recommendation E.180 [23]. Note that there are no specific guidelines for these tones. In the table, the symbol "+" indicates addition of the tones, without modulation, while "*" indicates amplitude modulation. The meaning of some of the tones is described in Section 3.12 or Section 3.11 (for V.21).

4.3 Use of RTP Header Fields

Timestamp: The RTP timestamp reflects the measurement point for the current packet. The event duration described in Section 4.4 extends forwards from that time.

4.4 Payload Format

Based on the characteristics described above, this document defines an RTP payload format called "tone" that can represent tones consisting of one or more frequencies. (The corresponding MIME type is "audio/tone".) The default timestamp rate is 8,000 Hz, but other rates may be defined. Note that the timestamp rate does not affect the interpretation of the frequency, just the durations.

Tone name	frequency	on period	off period
CNG	1100	0.5	3.0
V.25 CT	1300	0.5	2.0
CED	2100	3.3	–
ANS	2100	3.3	–
ANSam	2100*15	3.3	–
V.21 "0" bit, ch. 1	1180	0.00333	
V.21 "1" bit, ch. 1	980	0.00333	
V.21 "0" bit, ch. 2	1850	0.00333	
V.21 "1" bit, ch. 2	1650	0.00333	
ITU dial tone	425	–	–
U.S. dial tone	350+440	–	–
ITU ringing tone	425	0.67–1.5	3–5
U.S. ringing tone	440+480	2.0	4.0
ITU busy tone	425		
U.S. busy tone	480+620	0.5	0.5
ITU congestion tone	425		
U.S. congestion tone	480+620	0.25	0.25

Table 7: Examples of telephony tones

In accordance with current practice, this payload format does not have a static payload type number, but uses a RTP payload type number established dynamically and out-of-band.

It is shown in Fig. 3.

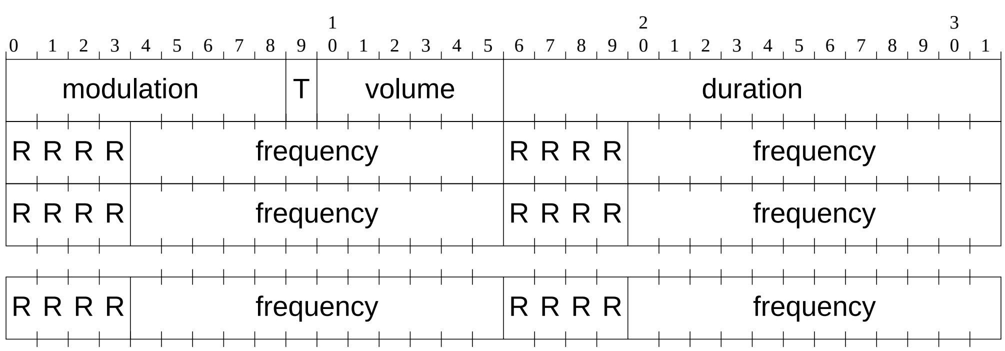

Figure 3: Payload format for tones

The payload contains the following fields:

modulation: The modulation frequency, in Hz. The field is a 9-bit unsigned integer, allowing modulation frequencies up to 511 Hz. If there is no modulation, this field has a value of zero.

T: If the "T" bit is set (one), the modulation frequency is to be divided by three. Otherwise, the modulation frequency is taken as is.

This bit allows frequencies accurate to 1/3 Hz, since modulation frequencies such as 16 2/3 Hz are in practical use.

volume: The power level of the tone, expressed in dBm0 after dropping the sign, with range from 0 to -63 dBm0. (Note: A preferred level range for digital tone generators is -8 dBm0 to -3 dBm0.)

duration: The duration of the tone, measured in timestamp units. The tone begins at the instant identified by the RTP timestamp and lasts for the duration value. The value of zero is not permitted and tones with such a duration SHOULD be ignored.

The definition of duration corresponds to that for sample-based codecs, where the timestamp represents the sampling point for the first sample.

frequency: The frequencies of the tones to be added, measured in Hz and represented as a 12-bit unsigned integer. The field size is sufficient to represent frequencies up to 4095 Hz, which exceeds the range of telephone systems. A value of zero indicates silence. A single tone can contain any number of frequencies.

R: This field is reserved for future use. The sender MUST set it to zero, the receiver MUST ignore it.

4.5 Reliability

This payload format uses the reliability mechanism described in Section 3.7.

5 Combining Tones and Named Events

The payload formats in Sections 3 and 4 can be combined into a single payload using the method specified in RFC 2198. Fig. 4 shows an example. In that example, the RTP packet combines two “tone” and one “telephone-event” payloads. The payload types are chosen arbitrarily as 97 and 98, respectively, with a sample rate of 8000 Hz. Here, the redundancy format has the dynamic payload type 96.

The packet represents a snapshot of U.S. ringing tone, 1.5 seconds (12,000 timestamp units) into the second “on” part of the 2.0/4.0 second cadence, i.e., a total of 7.5 seconds (60,000 timestamp units) into the ring cycle. The 440 + 480 Hz tone of this second cadence started at RTP timestamp 48,000. Four seconds of silence preceded it, but since RFC 2198 only has a fourteen-bit offset, only 2.05 seconds (16383 timestamp units) can be represented. Even though the tone sequence is not complete, the sender was able to determine that this is indeed ringback, and thus includes the corresponding named event.

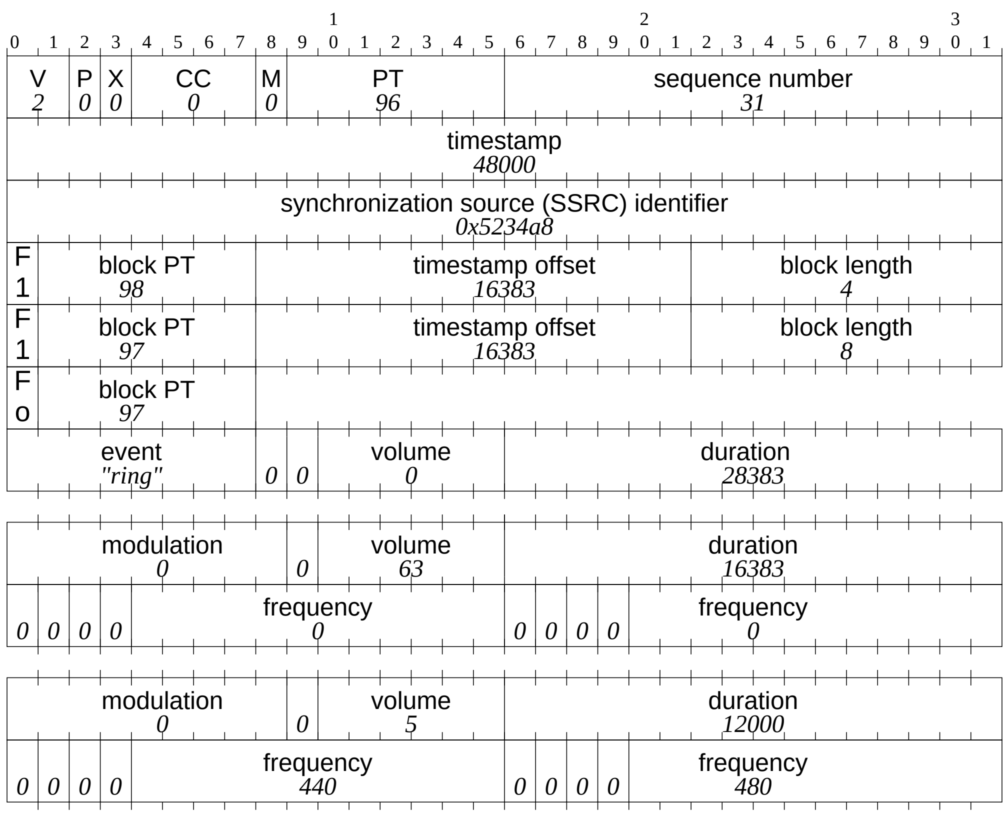

Figure 4: Combining tones and events in a single RTP packet

6 MIME Registration

6.1 audio/telephone-event

MIME media type name: audio

MIME subtype name: telephone-event

Required parameters: none.

Optional parameters: The "events" parameter lists the events supported by the implementation. Events are listed as one or more comma-separated elements. Each element can either be a single integer or two integers separated by a hyphen. No white space is allowed in the argument. The integers designate the event numbers supported by the implementation.

The "rate" parameter describes the sampling rate, in Hertz. The number is written as a floating point number or as an integer. If omitted, the default value is 8000 Hz.

Encoding considerations: This type is only defined for transfer via RTP [1].

Security considerations: See the "Security Considerations" (Section 7) section in this document.

Interoperability considerations: none

Published specification: This document.

Applications which use this media: The telephone-event audio subtype supports the transport of events occuring in telephone systems over the Internet.

Additional information: **1. Magic number(s):** N/A

2. File extension(s): N/A

3. Macintosh file type code: N/A

6.2 audio/tone

MIME media type name: audio

MIME subtype name: tone

Required parameters: none

Optional parameters: The "rate" parameter describes the sampling rate, in Hertz. The number is written as a floating point number or as an integer. If omitted, the default value is 8000 Hz.

Encoding considerations: This type is only defined for transfer via RTP [1].

Security considerations: See the "Security Considerations" (Section 7) section in this document.

Interoperability considerations: none

Published specification: This document.

Applications which use this media: The tone audio subtype supports the transport of pure composite tones, for example those commonly used in the current telephone system to signal call progress.

Additional information: **1. Magic number(s):** N/A

2. File extension(s): N/A

3. Macintosh file type code: N/A

7 Security Considerations

RTP packets using the payload format defined in this specification are subject to the security considerations discussed in the RTP specification (RFC 1889 [1]), and any appropriate RTP profile (for example RFC 1890 [24]).This implies that confidentiality of the media streams is achieved by encryption. Because the data compression used with this payload format is applied end-to-end, encryption may be performed after compression so there is no conflict between the two operations.

This payload type does not exhibit any significant non-uniformity in the receiver side computational complexity for packet processing to cause a potential denial-of-service threat.

In older networks employing in-band signaling and lacking appropriate tone filters, the tones in Section 3.14 may be used to commit toll fraud.

Additional security considerations are described in RFC 2198 [4].

8 IANA Considerations

This document defines two new RTP payload formats, named telephone-event and tone, and associated Internet media (MIME) types, audio/telephone-event and audio/tone.

Within the audio/telephone-event type, additional events MUST be registered with IANA. Registrations are subject to approval by the current chair of the IETF audio/video transport working group, or by an expert designated by the transport area director if the AVT group has closed.

The meaning of new events MUST be documented either as an RFC or an equivalent standards document produced by another standardization body, such as ITU-T.

9 Changes Since RFC 2833

- RFC 2833 had assigned only two code points to the three MF signals S1, S2 and S3. S3 has been moved to code point 174.
- The test tone descriptions were confusing; now, there are just two test tone entries, for the 2010 Hz and 1780 Hz tone.
- MFC R2 forward and backward tones were added to the trunk event list.
- Added the “trunk unavailable” event (Rajesh Kumar).
- Clarified that the duration timestamp is unsigned and that events exceeding the maximum duration expressible in the duration field should be split into several events, i.e., with a new start time.

- Distinguished states from events. States are sent with an estimated duration, and can be superseded if the state changes before the duration has expired. A special duration value of 0 indicates an infinite duration.
- Clarified how very long events that exceed the maximum expressable duration value should be handled.

10 Acknowledgements

The suggestions of the Megaco working group are gratefully acknowledged. Detailed advice and comments were provided by Hisham Abdelhamid, Flemming Andreasen, Fred Burg, Steve Casner, Dan Deliberato, Fatih Erdin, Bill Foster, Mike Fox, Mehryar Garakani, Gunnar Hellstrom, Rajesh Kumar, Terry Lyons, Steve Magnell, Zarko Markov, Kai Miao, Vern Paxson, Colin Perkins, Raghavendra Prabhu, Todd Sherer, Mira Stevanovic, Alex Urquizo and Herb Wildfeur.

11 Authors

Henning Schulzrinne
Dept. of Computer Science
Columbia University
1214 Amsterdam Avenue
New York, NY 10027
USA
electronic mail: schulzrinne@cs.columbia.edu

Scott Petrack
eDial
USA
electronic mail: scott.petrack@edial.com

Normative References

[1] H. Schulzrinne, S. Casner, R. Frederick, and V. Jacobson, “RTP: a transport protocol for real-time applications,” RFC 1889, Internet Engineering Task Force, Jan. 1996.

[2] S. Bradner, “Key words for use in RFCs to indicate requirement levels,” RFC 2119, Internet Engineering Task Force, Mar. 1997.

[3] International Telecommunication Union, “Procedures for starting sessions of data transmission over the public switched telephone network,” Recommendation V.8, Telecommunication Standardization Sector of ITU, Geneva, Switzerland, Feb. 1998.

[4] C. Perkins, I. Kouvelas, O. Hodson, V. Hardman, M. Handley, J. C. Bolot, A. Vega-Garcia, and S. Fosse-Parisis, “RTP payload for redundant audio data,” RFC 2198, Internet Engineering Task Force, Sept. 1997.

[5] International Telecommunication Union, “Multifrequency push-button signal reception,” Recommendation Q.24, Telecommunication Standardization Sector of ITU, Geneva, Switzerland, 1988.

[6] M. Handley and V. Jacobson, “SDP: session description protocol,” RFC 2327, Internet Engineering Task Force, Apr. 1998.

[7] International Telecommunication Union, “Automatic answering equipment and general procedures for automatic calling equipment on the general switched telephone network including procedures for disabling of echo control devices for both manually and automatically established calls,” Recommendation V.25, Telecommunication Standardization Sector of ITU, Geneva, Switzerland, Oct. 1996.

[8] International Telecommunication Union, “Procedures for document facsimile transmission in the general switched telephone network,” Recommendation T.30, Telecommunication Standardization Sector of ITU, Geneva, Switzerland, July 1996.

[9] International Telecommunication Union, “Echo cancellers,” Recommendation G.165, Telecommunication Standardization Sector of ITU, Geneva, Switzerland, Mar. 1993.

[10] International Telecommunication Union, “A modem operating at data signalling rates of up to 33 600 bit/s for use on the general switched telephone network and on leased point-to-point 2-wire telephone-type circuits,” Recommendation V.34, Telecommunication Standardization Sector of ITU, Geneva, Switzerland, Feb. 1998.

[11] International Telecommunication Union, “Procedures for the identification and selection of common modes of operation between data circuit-terminating equipments (DCEs) and between data terminal equipments (DTEs) over the public switched telephone network and on leased point-to-point telephone-type circuits,” Recommendation V.8bis, Telecommunication Standardization Sector of ITU, Geneva, Switzerland, Sept. 1998.

[12] International Telecommunication Union, “Operational and interworking requirements for DCEs operating in the text telephone mode,” Recommendation V.18, Telecommunication Standardization Sector of ITU, Geneva, Switzerland, Nov. 2000.

[13] International Telecommunication Union, “1200 bits per second duplex modem standardized for use in the general switched telephone network and on point-to-point 2-wire leased telephone-type circuits,” Recommendation V.22, Telecommunication Standardization Sector of ITU, Geneva, Switzerland, Nov. 1988.

[14] International Telecommunication Union, “Application of tones and recorded announcements in telephone services,” Recommendation E.182, Telecommunication Standardization Sector of ITU, Geneva, Switzerland, Mar. 1998.

[15] Bellcore, “Functional criteria for digital loop carrier systems,” Technical Requirement TR-NWT-000057, Telcordia (formerly Bellcore), Morristown, New Jersey, Jan. 1993.

[16] International Telecommunication Union, “Specifications of signalling system R2 – general,” Recommendation Q.440, Telecommunication Standardization Sector of ITU, Geneva, Switzerland, Nov. 1998.

[17] International Telecommunication Union, "Specifications for signaling system no. 5 – signal code for register signaling," Recommendation Q.151, Telecommunication Standardization Sector of ITU, Geneva, Switzerland, Nov. 1998.

[18] International Telecommunication Union, "AAL type 2 service specific convergence sublayer for trunking," Recommendation I.366.2, Telecommunication Standardization Sector of ITU, Geneva, Switzerland, Feb. 1999.

[19] International Telecommunication Union, "Specifications of signalling system R2 – forward line signals," Recommendation Q.400, Telecommunication Standardization Sector of ITU, Geneva, Switzerland, Nov. 1998.

[20] International Telecommunication Union, "Specifications of signalling system R2 – line signalling code," Recommendation Q.411, Telecommunication Standardization Sector of ITU, Geneva, Switzerland, Nov. 1998.

[21] International Telecommunication Union, "Specifications of signalling system R2 – digital line signalling code," Recommendation Q.421, Telecommunication Standardization Sector of ITU, Geneva, Switzerland, Nov. 1998.

[22] International Telecommunication Union, "Various tones used in national networks," Recommendation Supplement 2 to Recommendation E.180, Telecommunication Standardization Sector of ITU, Geneva, Switzerland, Jan. 1994.

[23] International Telecommunication Union, "Technical characteristics of tones for telephone service," Recommendation Supplement 2 to Recommendation E.180, Telecommunication Standardization Sector of ITU, Geneva, Switzerland, Jan. 1994.

[24] H. Schulzrinne, "RTP profile for audio and video conferences with minimal control," RFC 1890, Internet Engineering Task Force, Jan. 1996.

Informative References

[25] R. Kocen and T. Hatala, "Voice over frame relay implementation agreement," Implementation Agreement FRF.11, Frame Relay Forum, Foster City, California, Jan. 1997.

[26] J. G. van Bosse, *Signaling in Telecommunications Networks*. Telecommunications and Signal Processing, New York, New York: Wiley, 1998.

[27] Siemens, "MFC signaling systems," Jan. 1983. Siemens topics.