

AVT
Internet Draft
Intended status: Standards Track
Expires: April 2010

T. le Grand
Global IP Solutions
P. Jones
Cisco
P. Huart
Cisco
October 15, 2009

RTP Payload Format for the iSAC Codec
draft-legrand-rtp-isac-02.txt

Status of this Memo

This Internet-Draft is submitted to IETF in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

Copyright (c) 2009 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents in effect on the date of publication of this document (<http://trustee.ietf.org/license-info>). Please review these documents carefully, as they describe your rights and restrictions with respect to this document.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at <http://www.ietf.org/ietf/1id-abstracts.txt>

The list of Internet-Draft Shadow Directories can be accessed at <http://www.ietf.org/shadow.html>

This Internet-Draft will expire on April 15, 2010.

Abstract

iSAC is a proprietary wideband speech and audio codec developed by Global IP Solutions, suitable for use in Voice over IP applications. This document describes the payload format for iSAC generated bit streams within a Real-Time Protocol (RTP) packet. Also included here are the necessary details for the use of iSAC with the Session Description Protocol (SDP).

Conventions used in this document

In examples, "C:" and "S:" indicate lines sent by the client and server respectively.

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].

Table of Contents

1.	Introduction.....	2
2.	iSAC Codec Description.....	3
3.	RTP Payload Format.....	4
3.1.	iSAC Payload Format.....	4
3.2.	Payload Header.....	5
3.3.	Encoded Speech Data.....	5
3.4.	Multiple iSAC frames in an RTP packet.....	6
4.	IANA Considerations.....	6
4.1.	Media Type registration of iSAC.....	6
5.	Mapping to SDP Parameters.....	8
5.1.	Example Initial Target Bit Rate.....	8
5.2.	Example Max Bit Rate.....	9
6.	Security Considerations.....	9
7.	Acknowledgments.....	9
8.	References.....	9
8.1.	Normative References.....	9
8.2.	Informative References.....	10
	Author's Addresses.....	10

[1. Introduction](#)

This document gives a general description of the iSAC wideband speech codec and specifies the iSAC payload format for usage in RTP packets. Also included here are the necessary details for the use of iSAC with the Session Description Protocol (SDP).

2. iSAC Codec Description

The iSAC codec is an adaptive wideband speech and audio codec that operates with short delay, making it suitable for high quality real time communication. It is specially designed to deliver wideband speech quality in both low and medium bit rate applications. It also handles non-speech audio well, such as music and background noise [5].

The iSAC codec compresses speech frames of 16 kHz, 16-bit sampled input speech, each frame containing 30 or 60 ms of speech.

The codec runs in one of two different modes called channel-adaptive mode and channel-independent mode. In both modes iSAC is aiming at a target bit rate, which is neither the average nor the maximum bit rate that will be reached by iSAC, but corresponds to the average bit rate during peaks in speech activity. The bit rate will sometimes exceed the target bit rate, but most of the time will be below. The average bit rate obtained is on average about a factor of 1.4 times lower than the target bit rate.

In channel-adaptive mode the target bit rate is adapted to give a bit rate corresponding to the available bandwidth on the channel. The available bandwidth is constantly estimated at the receiving iSAC and signaled in-band in the iSAC bit stream. Even at dial-up modem data rates (including IP, UDP, and RTP overhead) iSAC delivers high quality by automatically adjusting transmission rates to give the best possible listening experience over the available bandwidth. The default initial target bit rate is 20000 bits per second in channel-adaptive mode.

In channel-independent mode a target bit rate has to be provided to iSAC prior to encoding.

After encoding the speech signal the iSAC coder uses lossless coding to further reduce the size of each packet, and hence the total bit rate used.

The adaptation and the lossless coding described above both result in a variation of packet size, depending both of the nature of speech and the available bandwidth. Therefore the iSAC codec operates at transmission rates from about 10 kbps to about 32 kbps.

The main characteristics can be summarized as follows:

- o Wideband, 16 kHz, speech and audio codec

- o Variable bit rate, which depends on the input signal
- o Adaptive rate with two modes: channel-adaptive or channel-independent mode
- o Bit rate range from around 10 kbps to 32 kbps
- o Operates on 30 or 60 ms of speech

3. RTP Payload Format

The iSAC codec uses a sampling rate clock of 16 kHz, so the RTP timestamp MUST be in units of 1/16000 of a second.

The RTP payload for iSAC has the format shown in Figure 1. No additional header fields specific to this payload format are required. For RTP based transportation of iSAC encoded audio, the standard RTP header [2] is followed by one payload data block.

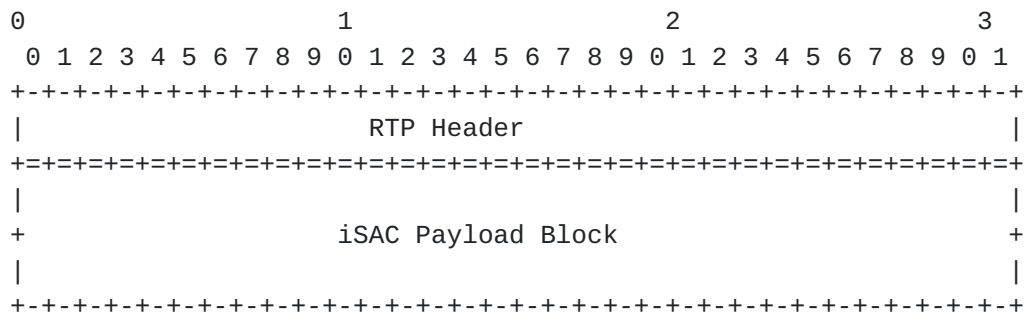


Figure 1: RTP packet format for iSAC

3.1. iSAC Payload Format

The iSAC payload block consists of a payload header and one or two encoded 30 ms speech frames. The iSAC payload is generated in the following manner:

- o Parameters representing one or two 30 ms frames of speech data are determined by the encoder. The parameters are quantized to generate encoded data corresponding to the one or two speech frames. The length of the encoded data is variable and depends on the signal characteristics and the target bit rate.
- o The payload header is generated (described in [Section 3.2](#)) and added before the encoded parameter data for the speech frame(s).

- o Lossless coding is applied to the complete iSAC payload block, including payload header, to generate a compressed payload. The length depends on the length of the data generated to represent the speech and the effectiveness of the lossless coding.

No part of the payload header or the encoded speech data can be retrieved without partly or fully decoding the packet.

The following figure shows an iSAC payload block containing 60 ms of encoded speech data:

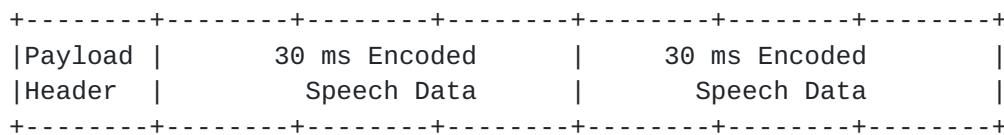


Figure 2: Payload format for iSAC

3.2. Payload Header

The payload header holds information for the receiver about the available bandwidth (BEI), and the length of the speech data in the current payload (FL). The header has the format defined in Figure 3. Note that the size of the header can vary due to the lossless encoding described in [section 2](#) and in [section 3.1](#). Also note that the BEI is always estimated and transmitted, even if iSAC runs in channel-independent mode.

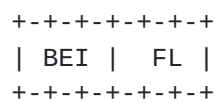


Figure 3: Payload Header

- o BEI: Bandwidth Estimation Index. The bandwidth estimate is quantized into one out of 24 values. Valid values are 0 to 23.
- o FL: The length of the speech data (Frame Length) present in the payload, given in number of speech samples. Valid frame lengths are 480 (30 ms) and 960 (60 ms) samples.

3.3. Encoded Speech Data

The iSAC encoded speech data consist of parameters representing one or two frames of 30 ms speech. The length of the speech data is signaled in the header (in number of samples), and the length may

change at any time during a session. In channel-adaptive mode the length is changed to best utilize the available bandwidth.

The iSAC payload is padded to whole octets, and has a variable length depending on the input source signal, number of 30 ms speech frames, and target bit rate.

The number of octets used to describe one frame of 30 ms speech typically varies from around 50 to around 120 octets. For the case of 60 ms speech (two 30 ms speech frames), the number of octets varies from around 100 to around 240 octets. The absolute maximum allowed payload length is 400 octets. The user can choose to lower the maximum allowed payload length. Minimum value is 100 octets. It is possible for the user to choose a maximum bit rate instead of a maximum payload length. The maximum payload length is then dependent on the length of the speech data represented in the payload (30 or 60 ms). Possible maximum rates are in the range of 32000 to 53400 bits per second.

The sensitivity to bit errors is equal for all bits in the payload.

3.4. Multiple iSAC frames in an RTP packet

More than one iSAC payload block MUST NOT be included in an RTP packet by a sender.

Further, iSAC payload blocks MUST NOT be split between RTP packets.

4. IANA Considerations

This document defines the iSAC media type.

4.1. Media Type registration of iSAC

Media type name: audio

Media subtype: isac

Required parameters: None

Optional parameters:

- o `ibitrate`: The parameter indicates the upper bound of the initial target bit rate the device would like to receive. For channel-adaptive mode, the target bit rate may vary with time; for channel-independent mode, the target bit rate will remain at that level unless instructed otherwise. An acceptable value for `ibitrate` is in the range of 20000 to 32000 (bits per second).
- o `maxbitrate`: The parameter indicates the maximum bit rate the endpoint expects to receive. The recipient of this parameter SHOULD NOT transmit at a higher bit rate.

Encoding considerations:

This media format is framed and binary.

Security considerations:

See [section 6](#).

Interoperability considerations: None

Published specification:

Applications which use this media type:

This media type is suitable for use in numerous applications needing to transport encoded voice or other audio. Some examples include Voice over IP, Streaming Media, Voice Messaging, and Conferencing.

Additional information: None

Intended usage: COMMON

Other Information/General Comment:

iSAC is a proprietary speech and audio codec owned by Global IP Solutions. The codec operates on 30 or 60 ms speech frames at a sampling rate clock of 16 kHz.

Person to contact for further information:

Tina le Grand [tina.legrand@gipscorp.com]

Restrictions on usage:

This media type depends on RTP framing, and hence is only defined for transfer via RTP [2]. Transport within other framing protocols is not defined at this time.

Change controller:

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

5. Mapping to SDP Parameters

The information carried in the media type specification has a specific mapping to fields in the Session Description Protocol (SDP) [4], which is commonly used to describe RTP sessions. When SDP is used to specify sessions employing the iSAC codec, the mapping is as follows:

- o The media type ("audio") goes in SDP "m=" as the media name.
- o The media subtype (payload format name) goes in SDP "a=rtpmap" as the encoding name.
- o Any remaining parameters go in the SDP "a=fmtp" attribute by copying them directly from the media type string as a semicolon separated list of parameter=value pairs.

The optional parameter `ibitrate` MUST NOT be higher than the parameter `maxbitrate`.

The iSAC parameters in an SDP offer are completely independent from those in the SDP answer. For both `ibitrate` and `maxbitrate` it is legal for the answer to contain a value that is different than what is provided in an offer. The parameter may be present in the answer, even if absent in the offer.

When conveying information by SDP, the encoding name SHALL be "isac" (the same as the media subtype).

5.1. Example Initial Target Bit Rate

The offer indicates that it wishes to receive a bitstream with an initial target rate of 20000 bits per second. The remote party MAY change its initial target rate to the requested value.

```
m=audio 10000 RTP/AVP 98
a=rtpmap: 98 isac/16000
a=fmtp:98 ibitrate=20000
```


5.2. Example Max Bit Rate

The offer indicates that it wishes to receive a bitstream with an initial target rate of 20000 bits per second, and a maximum bit rate of 45000 bits per second. The remote party MAY change its initial target rate and SHOULD NOT transmit at a higher rate than 45000.

```
m=audio 10000 RTP/AVP 98
a=rtpmap: 98 isac/16000
a=fmtp:98 ibitrate=20000;maxrate=45000
```

6. Security Considerations

RTP packets using the payload format defined in this specification are subject to the general security considerations discussed in [RFC 3550](#) [2].

As this format transports encoded speech, the main security issues include confidentiality and authentication of the speech itself. The payload format itself does not have any built-in security mechanisms. External mechanisms, such as SRTP [3], MAY be used.

7. Acknowledgments

This document was prepared using 2-Word-v2.0.template.dot.

8. References

8.1. Normative References

- [1] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), March 1997.
- [2] Schulzrinne, H., Casner, S., Frederick, R., and Jacobson, V., "RTP: A Transport Protocol for Real-Time Applications", STD 64, [RFC 3550](#), July 2003.
- [3] Baugher, M., McGrew, D., Naslund, M., Carrara, E., and Norrman, K., "The Secure Real-time Transport Protocol (SRTP)", [RFC 3711](#), March 2004.
- [4] Handley, M., Jacobson, V., and Perkins, C., "SDP: Session Description Protocol", [RFC 4566](#), July 2006.

8.2. Informative References

- [5] iSAC datasheet at Global IP Solutions website,
<http://www.gipscorp.com/files/english/datasheets/iSAC.pdf>

Author's Addresses

Tina le Grand
Global IP Solutions
Magnus Ladulasgatan 63B
SE-118 27 Stockholm
Sweden
Email: tina.legrand@gipscorp.com

Paul E. Jones
Cisco Systems, Inc,
7025 Kit Creek Rd.
Research Triangle Park, NC 27709
USA
Tel: +1 919 476 2048
Email: paulej@packetizer.com

Pascal Huart
Cisco Systems
400, Avenue Roumanille
Batiment T3
06410 BIOT - SOPHIA ANTIPOLIS
FRANCE
Tel: +33 4 9723 2643
Email: phuart@cisco.com