

Internet Engineering Task Force  
Internet draft

Gonzalo Camarillo  
Jan Heller  
Göran AP Eriksson  
Ericsson  
June 2000  
Expires March 2000

<mailto:camarillo@ericsson.com>  
<[draft-camarillo-sip-sdp-00.txt](#)>

## SDP media alignment in SIP

### Status of this Memo

This document is an Internet-Draft and is in full conformance with all provisions of [Section 10 of RFC2026](#).

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>.

### Abstract

This document defines an SDP media attribute. This attribute is intended to be used in conjunction with SIP in order to align different media streams belonging to a session. The use of this attribute allows sending media from a single media stream, encoded in different formats during the session, to different ports and host interfaces.

### 1. Introduction

SIP [[1](#)] is an application layer protocol for establishing, terminating and modifying multimedia sessions. SIP carries session descriptions in the bodies of the SIP messages but is independent from the protocol used for describing sessions. SDP [[2](#)] is one of the protocols that can be used for this purpose.

[Appendix B](#) of [[1](#)] describes the usage of SDP in relation to SIP. It

states: "The caller and callee align their media description so that the nth media stream ("m=" line) in the caller's session description corresponds to the nth media stream in the callee's description."

## SDP media alignment in SIP

This way of performing the media alignment implies that all the media from one single media stream has to be sent to the same port number. This does not suit fine systems that handle different codecs in different port numbers. Nor does it suit hosts supporting multiple interfaces that handle different codecs on different interface types.

### 2. SIP and cellular access

Systems using a cellular access (such as UMTS or EDGE) and SIP as a signalling protocol need to receive media over the air. During a session the media can be encoded using different codecs. The encoded media has to traverse the radio interface. The radio interface is generally characterized by being bit error prone and associated with relatively high packet transfer delays. In addition, radio interface resources in a cellular environment are scarce and thus expensive, which calls for special measures in providing a highly efficient transport [3]. In order to get an appropriate speech quality in combination with an efficient transport, precise knowledge of codec properties are required so that a proper radio bearer for the RTP session can be configured before transferring the media. These radio bearers are dedicated bearers per media type, i.e. codec.

In UMTS, for instance, when the RTP packets shall be delivered over the air interface, a packet filtering function routes the packets to the proper radio bearer towards the UMTS/SIP terminal. The packet filtering function operates using a Traffic Flow Template (TFT) [4], which is established when configuring the radio bearer. The TFT hence specifies the profile of the data that should be carried by the radio bearer. A TFT can contain the following data:

- Source Address and Subnet Mask.
- Protocol Number (IPv4) / Next Header (IPv6).
- Destination Port Range.
- Source Port Range.
- IPSec Security Parameter Index (SPI).
- Type of Service (TOS) (IPv4) / Traffic class (IPv6) and Mask.
- Flow Label (IPv6).

It is worth noticing that just certain combinations of these

parameters are allowed.

The media has to have different destination port numbers for the different possible codecs in order to be filtered and routed properly to the correct radio bearer. Examining the payload of every packet entering the UMTS network would require too much processing power. It would not be a viable solution. Demultiplexing of media streams should be done at the port level.

So, when a session is negotiated using SIP, a means for providing different port numbers for different codecs is needed. This procedure has to be backwards compatible with [appendix B](#) of [1].

### SDP media alignment in SIP

Systems using certain architectures have also this requirement in order to be able to use different audio libraries. Discussions about this can be found in [5].

### 3. SIP, SDP and multiple host interfaces

Two typical scenarios apply where a single media stream may be received or sent on different host interfaces. The first is when the cellular device is equipped with several access technologies. The second is the case where media format transcoding may be needed.

Cellular terminals are traditionally equipped with only one access type, the cellular access. Cellular terminals may also be further equipped with other higher speed wireless access technologies of short range nature such as Wireless LAN and Bluetooth.

Multiple accesses on the terminal implies that multiple IP addresses will be in use for a single terminal.

Applications based on SIP may rely on the fact that these additional interfaces may also be available for use, but of a more sporadic nature. The availability and applicability may for instance be based on coverage or simply by the user's preferences.

In a session invitation, the caller (on a mobile SIP terminal) may be willing to receive a single media stream that could be delivered on one of the available interfaces. A typical case could be where a stream of high quality audio can be supported over a high bit rate interface such as WLAN, and a low rate codec is used over the cellular access. Depending on the particular situation, the caller may offer both codec types, whereas the callee may select only one

of the codec types, e.g. depending on what he supports.

If the caller moves out of coverage of the high bit rate access, an ordinary re-invitation may move the media stream to be transported over the low bit rate cellular access.

The second scenario concerns the need for a transcoding service.

Speech codecs have been developed basically to be robust against the high bit error and frame error rate associated with wireless accesses. One such codec is GSM. Other encoding formats developed for wired accesses with negligible bit errors do not operate properly over a lossy link.

As fixed terminals may not support the wireless codecs, while cellular terminals can generally not support wired codecs, there is a need for a transcoding function between the two.

The transcoding function is typically placed in another host connected to the internet. This host is placed on the other side of the cellular access as viewed from the cellular terminal.

### SDP media alignment in SIP

To make use of the transcoding function, the cellular terminal may send an INVITE with an SDP part displaying support for the cellular codec as well as support for the codecs supported by the transcoder. In this situation, there is one media stream that depending on which codec is selected, should be directed to different host interfaces. The callee then makes a selection of codec and sends the media to the appropriate host interface. The transcoder will also be configured to forward the transcoded media stream to the cellular terminal.

#### [4.](#) Alignment of SDP in SIP to support multiple port numbers

A new "flow identification" media attribute is defined. It is used for identifying media streams within a session. It provides a means for aligning a number of flows within a session between members participating in the session. Its formatting in SDP is described by the following BNF:

```
fid-attribute      = "a=fid:" identification-tag
identification-tag = token
```

The identification tag is unique within the SDP session description. The following example illustrates its usage.

#### 5. Example of flow identification attribute

In the following example John uses a traditional access such as an ethernet while Laura has a UMTS/SIP terminal. The caller John sends the following session description to the callee Laura.

```
v=0
o=John 289085535 289085535 IN IP4 first.example.com
t=0 0
c=IN IP4 111.111.111.111
m=audio 20000 RTP/AVP 0 8
a=fid:1
m=audio 20002 RTP/AVP 0 8
a=fid:2
```

The callee Laura is on a UMTS/SIP terminal. She configures the necessary radio bearers and implements the TFTs:

For the first media stream with flow identification equal to 1 (fid:1):

All the incoming IP packets with destination port UDP 30000 will be carried by the radio access bearer configured for G-711 u-law (payload type 0).

All the incoming IP packets with destination port UDP 30002 will be carried by the radio access bearer configured for G-711 A-law (payload type 8).

#### SDP media alignment in SIP

For the second media stream with flow identification equal to 2 (fid:2):

All the incoming IP packets with destination port UDP 30004 will be carried by the radio access bearer configured for G-711 u-law (payload type 0).

All the incoming IP packets with destination port UDP 30006 will be carried by the radio access bearer configured for G-711 A-law (payload type 8).

Accordingly, the following SDP is returned to the caller:

```
v=0
o=Laura 289083124 289083124 IN IP4 second.example.com
t=0 0
c=IN IP4 222.222.222.222
m=audio 30000 RTP/AVP 0
a=fid:1
m=audio 30002 RTP/AVP 8
a=fid:1
m=audio 30004 RTP/AVP 0
a=fid:2
m=audio 30006 RTP/AVP 8
a=fid:2
```

In another example, the callee Laura might want to receive single flows in different IP addresses depending on the codec. It would return an SDP description as follows.

```
v=0
o=Laura 289083124 289083124 IN IP4 second.example.com
t=0 0
c=IN IP4 222.222.222.222
m=audio 30000 RTP/AVP 0
a=fid:1
m=audio 30002 RTP/AVP 8
c=IN IP4 123.123.123.123
a=fid:1
m=audio 30004 RTP/AVP 0
a=fid:2
m=audio 30006 RTP/AVP 8
c=IN IP4 123.123.123.123
a=fid:2
```

## [6. Media-level versus session-level attribute](#)

Syntactically `fid` is a media-level attribute. It provides information about a media stream defined by an "m" line. Semantically `fid` would be defined as a session-level attribute since it provides flow hierarchy inside a session description.

## [7. Backward compatibility](#)

Upon the reception of an SDP description with the `fid` attribute, the

callee must add also this attribute to the response. If the callee does not understand the fid attribute it will not include fid in the response. In this case the callee will proceed as it is described in [appendix B](#) of [1]. Thus, a flow will be created for every "m" line in the session description. The caller will have to issue a re-INVITE in order to remove the undesirable media streams. RFC2543bis prevents the caller to send an updated session description in an ACK request since a complete session description was sent already in the INVITE request. Until the re-INVITE is issued, the communication between the terminals is possible because they have enough information for receiving and sending media.

It would be possible to use different approaches that would resolve these issues (demultiplexing of media stream on port level) and provide better backward compatibility. One solution might be to add additional codecs supported in different IP addresses or port number in a new a attribute. This way, a callee that did not understand the attribute would ignore it but still create the correct number of media streams.

```
v=0
o=Laura 289083124 289083124 IN IP4 second.example.com
t=0 0
c=IN IP4 222.222.222.222
m=audio 30000 RTP/AVP 0
a=addcodec:8 222.222.222.222 30002
m=audio 30004 RTP/AVP 0
a=addcodec:8 222.222.222.222 30006
```

However, we believe that the fid attribute provides a more general mechanism to bundle different media lines into a single flow. This mechanism can be used by other applications different than the one described in this draft to overcome different problems. We have chosen generality over a specific solution.

The backward compatibility problem might not exist in a future depending on the format chosen for SDPng. This draft is also intended to serve as input for the SDPng work. Not on the syntax level but on the semantics of the fid attribute.

## [8](#). Open issues

How should RTCP be handled?

## [9](#). Acronyms

BNF	Backus-Naur Form
EDGE	Enhanced Data rates for GSM and TDMA/136 Evolution
GSM	Global System for Mobile communication

## SDP media alignment in SIP

IP	Internet Protocol
SDP	Session Description Protocol
SIP	Session Initiation Protocol
TFT	Traffic Flow Template
UMTS	Universal Mobile Telecommunication System
WLAN	Wireless Local Area Network

### 10. References

- [1] M. Handley/H. Schulzrinne/E. Schooler/J. Rosenberg, "SIP: Session Initiation Protocol", [RFC 2543](#), IETF; March 1999.
- [2] M. Handley/V. Jacobson, "SDP: Session Description Protocol", [RFC 2327](#), IETF; April 1998.
- [3] L. Westberg/M. Lindqvist, "Realtime Traffic over Cellular Access Networks", [draft-westberg-realtime-cellular-02.txt](#), IETF; May 2000. Work in progress.
- [4] 3G TS 23.060 v3.2.1 General Packet Radio Service Description.
- [5] <ftp://ftp.isi.edu/confctrl/confctrl.mail> ,thread with subject "SDP media alignment in SIP".

### 11. Authors' Addresses

Gonzalo Camarillo  
Ericsson  
Advanced Signalling Research Lab.  
FIN-02420 Jorvas  
Finland  
Phone: +358 9 299 3371  
Fax: +358 9 299 3118  
Email: Gonzalo.Camarillo@ericsson.com

Jan Höller  
Ericsson Research  
S-16480 Stockholm  
Sweden  
Phone: +46 8 58532845  
Fax: +46 8 4047020  
Email: Jan.Holler@era.ericsson.se

Göran AP Eriksson  
Ericsson Research  
S-16480 Stockholm

Sweden

Phone: +46 8 58531762

Fax: +46 8 4047020

Email: Goran.AP.Eriksson@era.ericsson.se

Camarillo/H÷ller/Eriksson

7