3GPP TSG CN Plenary Meeting #17 **NP-020360**
4th - 6th September 2002. Biarritz, France.

Title: **LS on Media grouping**
Source: **CN1**
Agenda item: **5.1**
Document for: **INFORMATION / ACTION**

3GPP TSG-CN1 Meeting #25 ***Tdoc N1-021782***
Helsinki, Finland, 29 July – 2 August

Title: LS on Media grouping
Response to:
Release: REL-5
Work Item: IMS-CCR

Source: CN1
To: SA, CN, SA2
Cc: CN3

Contact Person:
Name: Atle Monrad
Tel. Number: +47 372 93 665
E-mail Address: atle.monrad@ericsson.com

Attachments: N1-021675
draft-ietf-mmusic-fid-06.txt
draft-camarillo-mmusic-separate-streams-00.txt

1. Overall Description:

CN1 would like to inform about the status for the requirements described in 23.228 v.5.5.0 clause 4.2.5.1 (copied below).

4.2.5.1 Relation of IMS media components and PDP contexts carrying IMS media

The relation between IMS media components and PDP contexts carrying IMS media is controlled by the IMS network on media component level in the following way:
The P-CSCF shall have the capability to indicate to the UE that a separate PDP Context is required for each IMS media component indicated.

- If the UE receives such an indication for a media component, it shall open a separate PDP Context for this media component. If the UE receives no such indication for a media component, the UE makes the decision whether to open a separate PDP Context or modify an existing PDP Context for this media component.
- The criteria and information for setting this indication is determined by local policy in the network where the P-CSCF is located. The IMS network shall have the capability to transfer the media component level indication described above to the UE. This media component level indication shall be transferred in SIP/SDP signaling upon session initiation and addition of media component(s) to active IMS sessions.

It is assumed that media components from different IMS sessions are not carried within the same PDP context.
All associated IP flows (such as e.g. RTP / RTCP flows) used by the UE to support a single media component are assumed to be carried within the same PDP context.

From the referenced clause, it is the understanding of CN1 that separate media streams may be indicated from P-CSCF based on local policy. CN1 has evaluated the requirements and considered the following solutions:

1. Additions to SDP as described in draft-camarillo-mmusic-separate-streams-00.txt (additions to draft-ietf-mmusic-fid-06.txt currently close to being finalized in IETF).

2. Additions to SIP by e.g. a new p-header.

Both alternatives above require work in IETF, and CN1 has seen alternative 1 as the way forward and submitted draft-camarillo-mmusic-separate-streams-00.txt to IETF. The draft has got some comments but is not adopted as a working group item yet. The timeframe for this draft to possibly reach RFC status is not easy to predict, but it is the opinion of CN1 that the draft will at least need 6 months more to receive RFC status even if 3GPP puts effort into getting the document priority within IETF.

2. Actions:

To SA / CN groups.

ACTION:

CN1 asks SA / CN to consider the problem described above and come back to CN1 with guidance to the concerns for the Rel-5 timeframe.

Question 1.
It is a concern of CN1 that the functionality described above will cause a delay for finalising Rel-5 on time. Shall CN1 continue with the current working assumption and assume that draft-camarillo-mmusic-separate-streams-00.txt will reach RFC status in Rel-5 timeframe?

To SA2 group.

ACTION:

CN1 asks SA2 to consider the problem described above and come back to CN1 with further guidance to the questions below.

Question 2.
Does SA2 see other possible solutions to fulfil the requirement that does not cause additions to IETF and can be completed within the Rel-5 timeframe?

Question 3.
In case SA2 decides to move the requirement to indicate separate media streams as described in subclause 4.2.5.1 of 23.228 to Rel-6, CN1 would like to get guidance in how to proceed with this issue. Should CN1 continue the work as described in draft-camarillo-mmusic-separate-streams-00.txt for introduction in Rel-6?

3. Date of Next TSG-CN1 Meetings:

CN1_26	23rd – 27th September 2002	Miami, USA
CN1_27	11th – 15th November 2002	Bangkok, Thailand

Internet Engineering Task Force SIP WG
Internet Draft
G. Camarillo
Ericsson
A. Monrad
Ericsson
draft-camarillo-mmusic-separate-streams-00.txt
May 22, 2002
Expires: December 2002

Mapping of Media Streams to Resource Reservation Flows

STATUS OF THIS MEMO

Abstract

This document defines an extension to the SDP grouping framework. It allows to request that different media streams are mapped into different resource reservation flows.

Table of Contents

1 Introduction

Resource reservation protocols assign network resources to particular flows of IP packets. When a router receives an IP packet, it applies a filter in order to map the packet to the flow it belongs and provide it with the Quality of Service (QoS) corresponding to that flow. Routers typically use the source and the destination IP addresses and port numbers to filter packets.

Multimedia sessions typically contain multiple media streams (e.g. an audio stream and a video stream). In order to provide QoS for a multimedia session it is necessary to map all the media streams to resource reservation flows. This mapping can be performed in different ways. Two possibilities are to map all the media streams to a single resource reservation flow and to map every single media stream to a different resource reservation flow. Some applications require that the latter type of mapping is performed (i.e., a single media stream is mapped to a single resource reservation flow). This document defines the syntax needed to express that need in SDP [1]. For this purpose, we make use of the SDP grouping framework [2] and define a new "semantics" attribute called KIS (Keep It Separate).

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 [3] and indicate requirement levels for compliant SIP implementations.

2 KIS Semantics

We define a new "semantics" attribute within the SDP grouping framework [2]: KIS (Keep It Separate).

Media lines grouped using KIS semantics SHOULD NOT be mapped into the same resource reservation flow. A different resource reservation flow SHOULD be used (or established) for each media line of the KIS group.

3 Example

A user agent receives a SIP [4] INVITE with the SDP below:

```
v=0
o=Laura 289083124 289083124 IN IP4 one.example.com
t=0 0
c=IN IP4 192.0.0.1
```

```
a=group:KIS 1 2
m=audio 30000 RTP/AVP 0
a=mid:1
m=video 30002 RTP/AVP 31
a=mid:2
```

This user agent uses RSVP [5] to perform resource reservation. Since both media streams are part of a KIS group, the user agent will establish two different RSVP sessions; one for the audio stream and one for the video stream. An RSVP session is defined by the triple: (DestAddress, ProtocolId[, DstPort]). Table 1 shows the parameters used to establish both RSVP sessions.

Session Number	DestAddress	ProtocolId	DstPort
1	192.0.0.1	UDP	30000
2	192.0.0.1	UDP	30002

Table 1: Parameters needed to establish both RSVP sessions

If the same user agent received an SDP session description with the same media streams but without the group line, it would be free to map both media streams into the same RSVP session.

4 IANA Considerations

IANA needs to register the following new "semantics" attribute for the SDP grouping framework [2]:

KIS: Keep It Separate

5 Security Considerations

An attacker adding group lines using the KIS semantics to an SDP session description could force a user agent to establish a larger number of resource reservation flows than needed. It is thus RECOMMENDED that some kind of integrity protection is applied to SDP session descriptions.

6 Authors' Addresses

Gonzalo Camarillo
Ericsson
Advanced Signalling Research Lab.
FIN-02420 Jorvas
Finland
electronic mail: Gonzalo.Camarillo@ericsson.com

Atle Monrad
Ericsson
N-4898 Grimstad
Norway
electronic mail: atle.monrad@ericsson.com

7 Normative References

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

[2] G. Camarillo, J. Holler, G. Eriksson, and H. Schulzrinne, "Grouping of m lines in SDP," Internet Draft, Internet Engineering Task Force, Feb. 2002. Work in progress.

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

8 Informative References

[4] J. Rosenberg, H. Schulzrinne, et al. , "SIP: Session initiation protocol," Internet Draft, Internet Engineering Task Force, Feb. 2002. Work in progress.

[5] R. Braden, Ed., L. Zhang, S. Berson, S. Herzog, and S. Jamin, "Resource ReSerVation protocol (RSVP) -- version 1 functional specification," RFC 2205, Internet Engineering Task Force, Sept. 1997.

Full Copyright Statement

Internet Engineering Task Force
Internet draft

Gonzalo Camarillo
Jan Holler
Goran AP Eriksson
Ericsson

Henning Schulzrinne
Columbia University

February 2002
Expires August 2002
<draft-ietf-mmusic-fid-06.txt>

Grouping of media lines in SDP

Status of this Memo

Abstract

This document defines two SDP attributes: "group" and "mid". They allow to group together several "m" lines for two different purposes: for lip synchronization and for receiving media from a single flow (several media streams), encoded in different formats during a particular session, in different ports and host interfaces.

Grouping of media lines in SDP

TABLE OF CONTENTS

1 Introduction

An SDP session description typically contains a number (one or more) of media lines - they are commonly known as "m" lines. When a session description contains more than one "m" line, SDP does not provide any means to express a particular relationship between two or more of them. When an application receives an SDP session description with more than one "m" line it is up to the application what to do with them. SDP does not carry any information about grouping media streams.

While in some environments this information can be carried out of band, it would be desirable to have extensions to SDP that allowed to express how different media streams within a session description relate to each other. This document defines such extensions.

2 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 [1] and indicate requirement levels for compliant implementations.

3. Media stream identification attribute

A new "media stream identification" media attribute is defined. It is used for identifying media streams within a session description. Its formatting in SDP [2] is described by the following BNF:

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

The identification tag MUST be unique within an SDP session description.

4. Group attribute

A new "group" session level attribute is defined. It is used for grouping together different media streams. Its formatting in SDP is described by the following BNF:

```
group-attribute    = "a=group:" semantics
                       *(space identification-tag)
semantics          = "LS" | "FID"
```

This document defines two standard semantics: LS (Lip Synchronization) and FID (Flow Identification). If in the future it was needed to standardize further semantics they would need to be defined in a standards track document. However, defining new semantics apart from LS and FID is discouraged. Instead, it is RECOMMENDED to use other session description mechanisms such as SDPng.

5. Use of "group" and "mid"

All the "m" lines of a session description that uses "group" MUST be identified with an "mid" attribute whether they appear in the group line(s) or not. If a session description contains at least one "m" line that has no "mid" identification the application MUST NOT perform any grouping of media lines.

"a=group" lines are used to group together several "m" lines that are identified by their "mid" attribute. "a=group" lines that contain identification-tags that do not correspond to any "m" line within the session description MUST be simply ignored. The application acts as if the "a=group" line did not exist. The behavior of an application receiving an SDP with grouped "m" lines is defined by the semantics field in the "a=group" line.

There MAY be several "a=group" lines in a session description. All the "a=group" lines of a session description MAY or MAY NOT use the same semantics. An "m" line identified by its "mid" attribute MAY appear in more than one "a=group" line as long as the "a=group" lines use different semantics. An "m" line identified by its "mid" attribute MUST NOT appear in more than one "a=group" line using the same semantics.

An application that wants to be compliant to this specification MUST support both "group" and "mid". An application that supported just one of them would not be compliant.

6. Lip Synchronization (LS)

An application that receives a session description that contains "m" lines that are grouped together using LS semantics MUST synchronize the playout of the corresponding media streams. Note that LS semantics not only apply to a video stream that has to be synchronized with an audio stream. The playout of two streams of the same type can perfectly be synchronized as well.

For RTP streams synchronization is typically performed using RTCP, which provides enough information to map time stamps from the different streams into a wall clock. However, the concept of media stream synchronization MAY also apply to media streams that do not make use of RTP. If this is the case, the application MUST recover the original timing relationship between the streams using whatever available mechanism.

6.1 Example of LS

The following example shows a session description of a conference that is being multicast. The first media stream (mid:1) contains the voice of the speaker, who speaks in English. The second media stream (mid:2) contains the video component and the third (mid:3) media stream carries the translation to Spanish of what he is saying. The first and the second media streams MUST be synchronized.

```
v=0
o=Laura 289083124 289083124 IN IP4 one.example.com
t=0 0
c=IN IP4 224.2.17.12/127
a=group:LS 1 2
m=audio 30000 RTP/AVP 0
a=mid:1
m=video 30002 RTP/AVP 31
a=mid:2
m=audio 30004 RTP/AVP 0
i=This media stream contains the Spanish translation
a=mid:3
```

Note that although the third media stream is not present in the group line it still MUST contain an mid attribute (mid:3), as stated before.

7. Flow Identification (FID)

An "m" line in an SDP session description defines a media stream. However, SDP does not define what a media stream is. This definition can be found in the RTSP specification. The RTSP RFC [3] defines a media stream as "a single media instance, e.g., an audio stream or a video stream as well as a single whiteboard or shared application group. When using RTP, a stream consists of all RTP and RTCP packets created by a source within an RTP session".

This definition assumes that a single audio (or video) stream maps into an RTP session. The RTP RFC [4] defines an RTP session as follows: "For each participant, the session is defined by a particular pair of destination transport addresses (one network address plus a port pair for RTP and RTCP)".

While the previous definitions cover the most common cases, there are situations where a single media instance, (e.g., an audio stream or a video stream) is sent using more than one RTP session. Two examples (among many others) of this kind of situation are cellular systems using SIP [5] and systems receiving DTMF tones on a different host than the voice.

7.1 SIP and cellular access

Systems using a cellular access 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. 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.

Cellular systems typically configure different radio bearers on different port numbers. Therefore, incoming media has to have different destination port numbers for the different possible codecs in order to be routed properly to the correct radio bearer. Thus, this is an example in which several RTP sessions are used to carry a single media instance (the encoded speech from the sender).

7.2 DTMF tones

Some voice sessions include DTMF tones. Sometimes the voice handling is performed by a different host than the DTMF handling. It is common to have an application server in the network gathering DTMF tones for the user while the user receives the encoded speech on his user agent. In this situations it is necessary to establish two RTP sessions: one for the voice and the other for the DTMF tones. Both RTP sessions are logically part of the same media instance.

7.3 Media flow definition

The previous examples show that the definition of a media stream in [3] do not cover some scenarios. It cannot be assumed that a single media instance maps into a single RTP session. Therefore, we introduce the definition of a media flow:

Media flow consists of a single media instance, e.g., an audio stream or a video stream as well as a single whiteboard or shared application group. When using RTP, a media flow comprises one or more RTP sessions.

7.4 FID semantics

Several "m" lines grouped together using FID semantics form a media flow. A media agent handling a media flow that comprises several "m" lines MUST send a copy of the media to every "m" line part of the flow as long as the codecs and the direction attribute present in a particular "m" line allow it.

It is assumed that the application uses only one codec at a time to encode the media produced. This codec MAY change dynamically during the session, but at any certain moment only one codec is in use.

The application encodes the media using the current codec and checks one by one all the "m" lines that are part of the flow. If a particular "m" line contains the codec being used and the direction attribute is "sendonly" or "sendrecv" a copy of the encoded media is sent to the address/port specified in that particular media stream. If either the "m" line does not contain the codec being used or the direction attribute is neither "sendonly" nor "sendrecv" nothing is sent over this media stream.

The application typically ends up sending media to different destinations (IP address/port number) depending on the codec used at any moment.

7.4.1 Examples of FID

The session description below would be the SDP sent by a SIP user agent using a cellular access. The user agent supports GSM on port 30000 and AMR on port 30002. When the remote party sends GSM it will send RTP packets to port number 30000. When AMR is the codec chosen, packets will be sent to port 30002. Note that the remote party can switch between both codecs dynamically in the middle of the session. However, in this example, only one media stream at a time carries voice. The other remains "muted" while its corresponding codec is not in use.

```
      v=0
      o=Laura 289083124 289083124 IN IP4 two.example.com
      t=0 0
      c=IN IP4 131.160.1.112
      a=group:FID 1 2
      m=audio 30000 RTP/AVP 3
      a=rtpmap:3 GSM/8000
      a=mid:1
      m=audio 30002 RTP/AVP 97
      a=rtpmap:97 AMR/8000
      a=fmtp:97 mode-set=0,2,5,7; mode-change-period=2; mode-change-
   neighbor; maxframes=1
      a=mid:2
```

In the previous example a system receives media on the same IP address on different port numbers. The following example shows how a system can receive different codecs on different IP addresses.

```
      v=0
      o=Laura 289083124 289083124 IN IP4 three.example.com
      t=0 0
      c=IN IP4 131.160.1.112
      a=group:FID 1 2
      m=audio 20000 RTP/AVP 0
      c=IN IP4 131.160.1.111
      a=rtpmap:0 PCMU/8000
      a=mid:1
      m=audio 30002 RTP/AVP 97
      a=rtpmap:97 AMR/8000
      a=fmtp:97 mode-set=0,2,5,7; mode-change-period=2; mode-change-
   neighbor; maxframes=1
      a=mid:2
```

The cellular terminal of this example only supports the AMR codec. However, many current IP phones only support PCM (payload 0). In order to be able to interoperate with them, the cellular terminal uses a transcoder whose IP address is 131.160.1.111. The cellular terminal includes in its SDP support for PCM at that IP address. Remote systems will send AMR directly to the terminal but PCM will be sent to the transcoder. The transcoder will be configured (using whatever method) to convert the incoming PCM audio to AMR and send it to the terminal.

The next example shows that the "group" attribute used with FID semantics allows to express uni-directional codecs for a bi-directional media flow. That is, a codec that is only used in one direction within a sendrecv media stream.

```
v=0
o=Laura 289083124 289083124 IN IP4 four.example.com
t=0 0
c=IN IP4 131.160.1.112
a=group:FID 1 2
m=audio 30000 RTP/AVP 0
a=mid:1
m=audio 30002 RTP/AVP 8
a=recvonly
a=mid:2
```

A user agent that receives the SDP above knows that at a certain moment it can send either PCM u-law to port number 30000 or PCM A-law to port number 30002. However, the media agent also knows that the other end will only send PCM u-law (payload 0).

The following example shows a session description with different "m" lines grouped together using FID semantics that contain the same codec.

```
v=0
o=Laura 289083124 289083124 IN IP4 five.example.com
t=0 0
c=IN IP4 131.160.1.112
a=group:FID 1 2 3
m=audio 30000 RTP/AVP 0
a=mid:1
m=audio 30002 RTP/AVP 8
a=mid:2
m=audio 20000 RTP/AVP 0 8
c=IN IP4 131.160.1.111
a=recvonly
a=mid:3
```

At a particular point of time, if the media agent is sending PCM u-law (payload 0) it sends RTP packets to 131.160.1.112 on port 30000 and to 131.160.1.111 on port 20000 (first and third "m" lines). If it is sending PCM A-law (payload 8) it sends RTP packets to 131.160.1.112 on port 30002 and to 131.160.1.111 on port 20000 (second and third "m" lines).

The system that generated the SDP above supports PCM u-law on port 30000 and PCM A-law on port 30002. Besides, it uses an application server whose IP address is 131.160.1.111 that records all the conversation. That is why the application server always receives a copy of the audio stream regardless of the codec being used at any given moment (it actually performs an RTP dump, so it can effectively receive any codec).

Remember that if several "m" lines grouped together using FID semantics contain the same codec the media agent MUST send media over several RTP sessions at the same time.

The last example of this section deals with DTMF tones. DTMF tones can be transmitted using a regular voice codec or can be transmitted as telephony events. The RTP payload for DTMF tones treated as telephone events is described in RFC 2833 [6]. Below there is an example of an SDP session description using FID semantics and this payload type.

```
v=0
o=Laura 289083124 289083124 IN IP4 six.example.com
t=0 0
c=IN IP4 131.160.1.112
a=group:FID 1 2
m=audio 30000 RTP/AVP 0
a=mid:1
m=audio 20000 RTP/AVP 97
c=IN IP4 131.160.1.111
a=rtpmap:97 telephone-events
a=mid:2
```

The remote party would send PCM encoded voice (payload 0) to 131.160.1.112 and DTMF tones encoded as telephony events to 131.160.1.111. Note that only voice or DTMF is sent at a particular point of time. When DTMF tones are sent the first media stream does not carry any data and when voice is sent there is no data in the second media stream. FID semantics provide different destinations for alternative codecs.

7.5 Scenarios that FID does not cover

It is worthwhile mentioning some scenarios where the "group" attribute using existing semantics (particularly FID) might seem to be applicable but it is not. This section has been included because we have observed some confusion within the community regarding the three scenarios described below. This section helps clarify them.

7.5.1 Parallel encoding using different codecs

FID semantics are useful when the application only uses one codec at a time. An application that encodes the same media using different codecs simultaneously MUST NOT use FID to group those media lines. Some systems that handle DTMF tones are a typical example of parallel encoding using different codecs.

Some systems implement the RTP payload defined in RFC 2833, but when they send DTMF tones they do not mute the voice channel. Therefore, effectively they are sending two copies of the same DTMF tone: encoded as voice and encoded as a telephony event. When the receiver gets both copies it typically uses the telephony event rather than the tone encoded as voice. FID semantics MUST NOT be used in this context to group both media streams since such a system is not using alternative codecs but rather different parallel encodings for the same information.

7.5.2 Layered encoding

Layered encoding schemes encode media in different layers. Quality at the receiver varies depending on the number of layers received. SDP provides a means to group together contiguous multicast addresses that transport different layers. The "c" line below:

```
c=IN IP4 224.2.1.1/127/3
```

is equivalent to the following three "c" lines:

```
c=IN IP4 224.2.1.1/127
c=IN IP4 224.2.1.2/127
c=IN IP4 224.2.1.3/127
```

FID MUST NOT be used to group "m" lines that do not represent the same information. Therefore, FID MUST NOT be used to group "m" lines that contain the different layers of layered encoding scheme. Besides, we do not define new group semantics to provide a more flexible way of grouping different layers because the already existing SDP mechanism covers the most useful scenarios.

7.5.3 Same IP address and port number

If several codecs have to be sent to the same IP address and port, the traditional SDP syntax of listing several codecs in the same "m" line MUST be used. FID MUST NOT be used to group "m" lines with the same IP address/port. Therefore, an SDP like the one below MUST NOT be generated.

```
v=0
o=Laura 289083124 289083124 IN IP4 six.example.com
t=0 0
c=IN IP4 131.160.1.112
a=group:FID 1 2
m=audio 30000 RTP/AVP 0
a=mid:1
m=audio 30000 RTP/AVP 8
a=mid:2
```

The correct SDP for the session above would be the following one:

```
v=0
o=Laura 289083124 289083124 IN IP4 six.example.com
t=0 0
c=IN IP4 131.160.1.112
m=audio 30000 RTP/AVP 0 8
```

If two "m" lines are grouped using FID they MUST differ in their transport addresses (i.e., IP address plus port).

8. Usage of the "group" attribute in SIP

SDP descriptions are used by several different protocols, SIP among them. We include a section about SIP because the "group" attribute will most likely be used mainly by SIP systems.

SIP [5] 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.

At session establishment SIP provides a three-way handshake (INVITE-200 OK-ACK) between end systems. However, just two of these three messages carry SDP. SDPs MAY be present in INVITE and 200 OK or in 200 OK and ACK. The following sections assume that INVITE and 200 OK are the ones carrying SDP for the sake of clarity, but everything is also applicable to the other possible scenario (200 OK and ACK).

8.1 Mid value in responses

The "mid" attribute is an identifier for a particular media stream. Therefore, the "mid" value in the response MUST be the same as the "mid" value in the request. Besides, subsequent requests such as re-INVITEs SHOULD use the same "mid" value for the already existing media streams.

Appendix B of [5] 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 caller³s session description corresponds to the nth media stream in the callee³s description."

The presence of the "group" attribute in an SDP session description does not modify this behavior.

Since the "mid" attribute provides a means to label "m" lines it would be possible to perform media alignment using "mid" labels rather than matching nth "m" lines. However this would not bring any gain and would add complexity to implementations. Therefore SIP systems MUST perform media alignment matching nth lines regardless of the presence of the "group" or "mid" attributes.

If a media stream that contained a particular "mid" identifier in the request contains a different identifier in the response the application ignores all the "mid" and "group" lines that might appear in the session description. The following example illustrates this scenario:

8.1.1 Example

Two SIP entities exchange SDPs during session establishment. The INVITE contained the SDP below:

```
v=0
o=Laura 289083124 289083124 IN IP4 seven.example.com
t=0 0
c=IN IP4 131.160.1.112
a=group:FID 1 2
m=audio 30000 RTP/AVP 0 8
a=mid:1
m=audio 30002 RTP/AVP 0 8
a=mid:2
```

The 200 OK response contains the following SDP:

```
v=0
o=Bob 289083122 289083122 IN IP4 eigth.example.com
t=0 0
c=IN IP4 131.160.1.113
a=group:FID 1 2
m=audio 25000 RTP/AVP 0 8
a=mid:2
m=audio 25002 RTP/AVP 0 8
a=mid:1
```

Since alignment of "m" lines is performed based on matching of nth lines, the first stream had "mid:1" in the INVITE and "mid:2" in the 200 OK. Therefore, the application MUST ignore every "mid" and "group" lines contained in the SDP.

A well-behaved SIP user agent would have returned the SDP below in the 200 OK:

```
v=0
o=Bob 289083122 289083122 IN IP4 nine.example.com
t=0 0
c=IN IP4 131.160.1.113
a=group:FID 1 2
m=audio 25002 RTP/AVP 0 8
a=mid:1
m=audio 25000 RTP/AVP 0 8
a=mid:2
```

8.2 Group value in responses

A SIP entity that receives a request that contains an "a=group" line with semantics that it does not understand MUST return a response without the "group" line. Note that, as it was described in the previous section, the "mid" lines MUST still be present in the response.

A SIP entity that receives a request that contains an "a=group" line which semantics that are understood MUST return a response that contains an "a=group" line with the same semantics. The identification-tags contained in this "a=group" lines MUST be the same that were received in the request or a subset of them (zero identification-tags is a valid subset). When the identification-tags in the response are a subset the "group" value to be used in the session MUST be the one present in the response.

SIP entities refuse media streams by setting the port to zero in the corresponding "m" line. "a=group" lines MUST NOT contain identification-tags that correspond to "m" lines with port zero.

Note that grouping of m lines MUST always be requested by the issuer of the request (the client), never by the issuer of the response (the server). Since SIP provides a two-way SDP exchange, a server that requested grouping in a response would not know whether the "group" attribute was accepted by the client or not. A server that wants to group media lines SHOULD issue another request after having responded to the first one (a re-INVITE for instance).

> Note that, as we mentioned previously, in this section we are assuming that the SDPs are present in the INVITE and in the 200 OK. Applying the statement above to the scenario where SDPs are present in the 200 OK and in the ACK, the entity requesting grouping would be the server.

8.2.1 Example

The example below shows how the callee refuses a media stream offered by the caller by setting its port number to zero. The "mid" value corresponding to that media stream is removed from the "group" value in the response.

SDP in the INVITE from caller to callee:

```
v=0
o=Laura 289083124 289083124 IN IP4 ten.example.com
t=0 0
c=IN IP4 131.160.1.112
a=group:FID 1 2 3
m=audio 30000 RTP/AVP 0
a=mid:1
m=audio 30002 RTP/AVP 8
a=mid:2
m=audio 30004 RTP/AVP 3
a=mid:3
```

SDP in the INVITE from callee to caller:

```
v=0
```

```
o=Bob 289083125 289083125 IN IP4 eleven.example.com
```

```
t=0 0
c=IN IP4 131.160.1.113
a=group:FID 1 3
m=audio 20000 RTP/AVP 0
a=mid:1
m=audio 0 RTP/AVP 8
a=mid:2
m=audio 20002 RTP/AVP 3
a=mid:3
```

8.3 Capability negotiation

A client that understands "group" and "mid" but does not want to make use of them in a particular session MAY want indicate that it supports them. If a client decides to do that, it SHOULD add an "a=group" line with zero identification-tags for every semantics it understands.

If a server receives a request that contains empty "a=group" lines it SHOULD add its capabilities also in the form of empty "a=group" lines to its response.

8.3.1 Example

A system that supports both LS and FID semantics but does not want to group any media stream for this particular session generates the following SDP:

```
v=0
o=Bob 289083125 289083125 IN IP4 twelve.example.com
t=0 0
c=IN IP4 131.160.1.113
a=group:LS
a=group:FID
m=audio 20000 RTP/AVP 0 8
```

The server that receives that request supports FID but not LS. It responds with the SDP below:

```
v=0
o=Laura 289083124 289083124 IN IP4 thirteen.example.com
t=0 0
c=IN IP4 131.160.1.112
a=group:FID
m=audio 30000 RTP/AVP 0
```

8.4 Backward compatibility

This document does not define any SIP "Require" header. Therefore, if one of the SIP user agents does not understand the "group"

attribute the standard SDP fall back mechanism MUST be used (attributes that are not understood are simply ignored).

8.4.1 Client does not support "group"

This situation does not represent a problem because grouping requests is always performed by clients, not by servers. If the client does not support "group" this attribute will just not be used.

8.4.2 Server does not support "group"

The server will ignore the "group" attribute, since it does not understand it (it will also ignore the "mid" attribute). For LS semantics, the server might decide to perform or to not perform synchronization between media streams.

For FID semantics, the server will consider that the session comprises several media streams.

Different implementations would behave in different ways.

In the case of audio and different "m" lines for different codecs an implementation might decide to act as a mixer with the different incoming RTP sessions, which is the correct behavior.

An implementation might also decide to refuse the request (e.g. 488 Not acceptable here or 606 Not Acceptable) because it contains several "m" lines. In this case, the server does not support the type of session that the caller wanted to establish. In case the client is willing to establish a simpler session anyway, he SHOULD re-try the request without "group" attribute and only one "m" line per flow.

9. Security considerations

Using the "group" parameter with FID semantics an entity that managed to modify the session descriptions exchanged between the participants to establish a multimedia session could force the participants to send a copy of the media to any particular destination.

Integrity mechanism provided by protocols used to exchange session descriptions and media encryption can be used to prevent this attack.

10. IANA considerations

This document defines two SDP attributes: "mid" and "group".

The "mid" attribute is used to identify media streams within a session description and its format is defined in Section 3.

The "group" attribute is used for grouping together different media streams and its format is defined in Section 4.

Section 4 also defines two standard semantics related to the "group" attribute: LS (Lip Synchronization) and FID (Flow Identification). If in the future it was needed to standardize further semantics they would need to be defined in a standards track document.

11. Acknowledgments

The authors would like to thank Jonathan Rosenberg, Adam Roach, Orit Levin and Joerg Ott for their feedback on this document.

12. References

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

[2] M. Handley/V. Jacobson, "SDP: Session Description Protocol", RFC 2327, IETF; April 1998.

[3] H. Schulzrinne/A. Rao/R. Lanphier, "Real Time Streaming Protocol (RTSP)", RFC 2326, IETF; April 1998.

[4] H. Schulzrinne/S. Casner/R. Frederick/V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications", RFC 1889, IETF; January 1996.

[5] M. Handley/H. Schulzrinne/E. Schooler/J. Rosenberg, "SIP: Session Initiation Protocol", RFC 2543, IETF; Mach 1999.

[6] H. Schulzrinne/S. Petrack, "RTP Payload for DTMF Digits, Telephony Tones and Telephony Signals", RFC 2833, IETF; May 2000.

13. Authors³ Addresses

Gonzalo Camarillo
Ericsson
Advanced Signalling Research Lab.
FIN-02420 Jorvas
Finland
Phone: +358 9 299 3371
Fax: +358 9 299 3052

Email: Gonzalo.Camarillo@ericsson.com

Grouping of media lines in SDP

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

Goran AP Eriksson
Ericsson Research
S-16480 Stockholm
Sweden
Phone: +46 8 58531762
Fax: +46 8 4047020
Email: Goran.AP.Eriksson@era.ericsson.se

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

Full Copyright Statement

Grouping of media lines in SDP

3GPP TSG-CN1 Meeting #25
Helsinki, Finland, 29 July – 2 August

Tdoc N1-021675

CR-Form-v7

CHANGE REQUEST

⌘	24.229	CR	175	⌘ rev	-	⌘	Current version:	5.1.0	⌘

*For **HELP** on using this form, see bottom of this page or look at the pop-up text over the ⌘ symbols.*

Proposed change affects:	UICC apps⌘		ME	X	Radio Access Network		Core Network

Title:	⌘	Clarifications of the binding and media grouping	
Source:	⌘	Ericsson	
Work item code:	⌘	IMS-CCR	***Date:*** ⌘ 22/07/2002
Category:	⌘	**F**	***Release:*** ⌘ REL-5
		Use one of the following categories: ***F*** *(correction)* ***A*** *(corresponds to a correction in an earlier release)* ***B*** *(addition of feature),* ***C*** *(functional modification of feature)* ***D*** *(editorial modification)* Detailed explanations of the above categories can be found in 3GPP TR 21.900.	*Use one of the following releases:* *2 (GSM Phase 2)* *R96 (Release 1996)* *R97 (Release 1997)* *R98 (Release 1998)* *R99 (Release 1999)* *Rel-4 (Release 4)* *Rel-5 (Release 5)* *Rel-6 (Release 6)*

Reason for change:	⌘	Reference to the internet drafts providing the detailed working assumption for the grouping of m-lines replace the reference of the stage-2 reference. Various clarifications of the text describing binding.
Summary of change:	⌘	Clarifications when and how a separate PDP context is required for media is needed in order to get consistant behaviour of the UE from different vendors.
Consequences if not approved:	⌘	The detailed description of binding and grouping information will not be complete.

Clauses affected:	⌘	2, 9.2.5				
		Y	N			
Other specs Affected:	⌘	X		Other core specifications	⌘	27.060
			X	Test specifications		
			X	O&M Specifications		
Other comments:	⌘					

How to create CRs using this form:
Comprehensive information and tips about how to create CRs can be found at http://www.3gpp.org/specs/CR.htm. Below is a brief summary:

1) Fill out the above form. The symbols above marked ⌘ contain pop-up help information about the field that they are closest to.

2) Obtain the latest version for the release of the specification to which the change is proposed. Use the MS Word "revision marks" feature (also known as "track changes") when making the changes. All 3GPP specifications can be downloaded from the 3GPP server under ftp://ftp.3gpp.org/specs/ For the latest version, look for the directory name with the latest date e.g. 2001-03 contains the specifications resulting from the March 2001 TSG meetings.

3) With "track changes" disabled, paste the entire CR form (use CTRL-A to select it) into the specification just in front of the clause containing the first piece of changed text. Delete those parts of the specification which are not relevant to the change request.

**************** First modification ********************

2 References

The following documents contain provisions which, through reference in this text, constitute provisions of the present document.

- References are either specific (identified by date of publication, edition number, version number, etc.) or non-specific.
- For a specific reference, subsequent revisions do not apply.
- For a non-specific reference, the latest version applies. In the case of a reference to a 3GPP document (including a GSM document), a non-specific reference implicitly refers to the latest version of that document *in the same Release as the present document*.

[1] 3GPP TR 21.905: "Vocabulary for 3GPP Specifications".

[2] 3GPP TS 23.002: "Network architecture".

[3] 3GPP TS 23.003: "Numbering, addressing and identification".

[4] 3GPP TS 23.060: "General Packet Radio Service (GPRS); Service description; Stage 2".

[5] 3GPP TS 23.218: "IP Multimedia (IM) Session Handling; IM call model".

[6] 3GPP TS 23.221: "Architectural requirements".

[7] 3GPP TS 23.228: "IP multimedia subsystem; Stage 2".

[8] 3GPP TS 24.008: "Mobile radio interface layer 3 specification; Core Network protocols; Stage 3".

[9] 3GPP TS 25.304: "UE Procedures in Idle Mode and Procedures for Cell Reselection in Connected Mode".

[10] 3GPP TS 26.235: "Packet switched conversational multimedia applications; Default codecs".

[11] 3GPP TS 29.061: "Interworking between the Public Land Mobile Network (PLMN) supporting Packet Based Services and Packet Data Networks (PDN)".

[12] 3GPP TS 29.207: "Policy control over Go interface".

[13] 3GPP TS 29.208: "End to end Quality of Service (QoS) signalling flows".

[14] 3GPP TS 29.228: "IP Multimedia (IM) Subsystem Cx and Dx Interfaces; Signalling flows and message contents".

[15] 3GPP TS 29.229: "Cx and Dx Interfaces based on the Diameter protocol, Protocol details".

[16] 3GPP TS 32.200: "Telecommunication management; Charging management; Charging principles".

[17] 3GPP TS 32.225: "Telecommunication management; Charging management; Charging data description for the IP Multimedia subsystem".

[18] 3GPP TS 33.102: "3G Security; Security architecture".

[19] 3GPP TS 33.203: "Access security for IP based services".

[20] 3GPP TS 44.018: "Mobile radio interface layer 3 specification, Radio Resource Control Protocol".

[21] RFC 2617: "HTTP Authentication: Basic and Digest Access Authentication".

[22] RFC 2806: "URLs for Telephone Calls".

[23] RFC 2833 (May 2000): "RTP Payload for DTMF Digits, Telephony Tones and Telephony Signals".

[24] RFC 2916: "E.164 number and DNS".

[25] RFC 2976 (October 2000): "The SIP INFO method".

[26] RFC 3261 (March 2002): "SIP: Session Initiation Protocol".

[27] RFC 3262 (March 2002): "Reliability of provisional responses in Session Initiation Protocol".

[28] RFC 3265 (March 2002): "Session Initiation Protocol Specific Event Notification".

[29] RFC 3311 (April 2002): "The SIP UPDATE method".

[30] RFC 3312 (May 2002): "Integration of resource management and SIP".

[31] RFC 3313 (February 2002): "SIP extensions for media authorization".

[32] RFC 3320 (March 2002): "Signaling Compression (SigComp)"

[33] RFC 3323 (May 2002): "A Privacy Mechanism for the Session Initiation Protocol (SIP)".

[34] RFC 3325 (May 2002): "Private Extensions to the Session Initiation Protocol (SIP) for Network Asserted Identity within Trusted Networks".

[35] RFC 3327 (May 2002): "SIP Extension for Registering Non-Adjacent Contacts".

[36] draft-sparks-sip-refer-split-00 (April 2002): "The REFER method".

Editor's note: The above document cannot be formally referenced until it is published as an RFC.

[37] draft-sparks-sip-mimetypes (April 2002): "Internet Media Type message/sipfrag".

Editor's note: The above document cannot be formally referenced until it is published as an RFC.

[38] draft-willis-scvrtdisco-03 (May 2002): "SIP Extension Header for Service Route Discovery in Private Networks".

Editor's note: The above document cannot be formally referenced until it is published as an RFC.

[39] draft-ietf-mmusic-sdp-new-04 (November 2001): "SDP: Session Description Protocol".

Editor's note: The above document cannot be formally referenced until it is published as an RFC.

[40] draft-ietf-dhc-dhcpv6-23 (February 2002): "Dynamic Host Configuration Protocol for IPv6 (DHCPv6)".

Editor's note: The above document cannot be formally referenced until it is published as an RFC.

[41] draft-iet f-sip-dhcpv6-00 (April 2002): "DHCPv6 options for SIP servers".

Editor's note: The above document cannot be formally referenced until it is published as an RFC.

[42] draft-ietf-sipping-sigcomp-sip-dictionary-00.txt (May 2002): "The SIP/SDP static dictionary for Signaling Compression".

Editor's note: The above document cannot be formally referenced until it is published as an RFC.

[43] draft-beckmann-sip-reg-event-01 (May 2002): "Registration event package".

Editor's note: The above document cannot be formally referenced until it is published as an RFC.

[44] draft-garcia-sip-visited-network-id-00 (March 2002): "Private SIP extension for Visited Network Identifier".

Editor's note: The above document cannot be formally referenced until it is published as an RFC.

[45] draft-henrikson-sip-charging-information-01 (May 2002): "Private SIP Extension for Mobile Charging Information".

Editor's note: The above document cannot be formally referenced until it is published as an RFC.

[46] draft-henrikson-sip-original-dialog-id-01 (May 2002): "Private SIP Extension for Original Dialog Identifier".

Editor's note: The above document cannot be formally referenced until it is published as an RFC.

[47] draft-mills-sip-access-network-info-01.txt (April 2002): "~~“~~SIP Access Network Information header~~”~~".

Editor's note: The above document cannot be formally referenced until it is published as an RFC.

[48] draft-ietf-mmusic-fid-06 (August 2002): "Grouping of media lines in SDP".

Editor's note: The above document cannot be formally referenced until it is published as an RFC.

[49] draft-camarillo-mmusic-separate-streams-00 (December 2002): "Mapping of Media Streams to Resource Reservation Flows".

Editor's note: The above document cannot be formally referenced until it is published as an RFC.

*************** Second modification ********************

9 GPRS aspects when connected to the IM CN subsystem

9.1 Introduction

A UE accessing the IM CN subsystem, and the IM CN subsystem itself, utilise the services provided by GPRS to provide packet-mode communication between the UE and the IM CN subsystem.

Requirements for the UE on the use of these packet-mode services are specified in this clause. Requirements for the GGSN in support of this communication are specified in 3GPP TS 29.061 [11] and 3GPP TS 29.207 [12].

9.2 Procedures at the UE

9.2.1 PDP context activation and P-CSCF discovery

Prior to communication with the IM CN subsystem, the UE shall:

a) perform a GPRS attach procedure;

b) establish a PDP context used for SIP signalling according to the APN and GGSN selection criteria described in 3GPP TS 23.060 [4]. This PDP context shall remain active throughout the period the UE is connected to the IM CN subsystem, i.e. from the initial registration and at least until the deregistration. As a result, the PDP context provides the UE with information that makes the UE able to construct an IPv6 address;

   The UE shall choose one of the following options when performing establishment of this PDP context:

   I. A dedicated PDP context for SIP signalling:

      The UE shall indicate to the GGSN that this is a PDP context intended to carry IM CN subsystem-related signalling only by setting the IM CN Subsystem Signalling Flag within the Protocol Configuration Options IE at PDP Context activation. The UE may also use this PDP context for DNS and DHCP signalling according to the static packet filters described in 3GPP TS 29.207 [12];

II. A general-purpose PDP context:

The UE may decide to use a general purpose PDP Context to carry IM CN subsystem-related signaling. The UE shall indicate to the GGSN that this is a general-purpose PDP context by not setting the IM CN Subsystem Signalling Flag within the Protocol Configuration Options IE;

NOTE 1: A general purpose PDP Context is completely IM CN subsystem-unaware, and as such, it does not have any IM CN subsystem-specific mechanisms applied to it.

NOTE 2: A general purpose PDP Context may carry both IM CN subsystem signaling and media, in case the media does not need to be authorized by Service Based Local Policy mechanisms defined in 3GPP TS 29.207 [12] and the media component is not mandated by the P-CSCF to be carried in a separate PDP Context.

c) aquire a P-CSCF address(es).

The methods for P-CSCF discovery are:

I. Employ Dynamic Host Configuration Protocol for IPv6 (DHCPv6) draft-ietf-dhc-dhcpv6 [40], the DHCPv6 options for SIP servers draft-ietf-sip-dhcpv6 [41] and if needed DNS after PDP context activation.

The UE shall either:

- in the DHCP query, request a list of SIP server domain names of P-CSCF(s) and the list of Domain Name Servers (DNS); or

- request a list of SIP server IPv6 addresses of P-CSCF(s).

II. Transfer P-CSCF address(es) within The PDP context activation procedure.

The UE shall indicate the request for a P-CSCF address to the GGSN within the Protocol Configuration Options IE of the ACTIVATE PDP CONTEXT REQUEST message or ACTIVATE SECONDARY PDP CONTEXT REQUEST message.

If the GGSN provides the UE with a list of P-CSCF IPv6 addresses in the ACTIVATE PDP CONTEXT ACCEPT message or ACTIVATE SECONDARY PDP CONTEXT ACCEPT message, the UE shall assume that the list is prioritised with the first address within the Protocol Configuration Options IE as the P-CSCF address with the highest priority.

The UE can freely select method I or II for P-CSCF discovery. In case several P-CSCF addresses are provided to the UE, the selection of P-CSCF address shall be performed according to the resolution of host name as indicated in RFC 3261 [26]. If sufficient information for P-CSCF address selection is not available, selection of the P-CSCF address by the UE is implementation specific.

If the UE is designed to use I above, but receives P-CSCF address(es) according to II, then the UE shall either ignore the received address(es), or use the address(es) in accordance with II, and not proceed with the DHCP request according to I.

9.2.2 Session management procedures

The existing procedures for session management as described in 3GPP TS 24.008 [8] shall apply while the UE is connected to the IM CN subsystem.

9.2.3 Mobility management procedures

The existing procedures for mobility management as described in 3GPP TS 24.008 [8] shall apply while the UE is connected to the IM CN subsystem.

9.2.4 Cell selection and lack of coverage

The existing mechanisms and criteria for cell selection as described in 3GPP TS 25.304 [9] and 3GPP TS 44.018 [20] shall apply while the UE is connected to the IM CN subsystem.

9.2.5 PDP contexts for media

During establishment of a session, the UE establishes data streams(s) for media related to the session. Such data stream(s) may result in activation of additional PDP context(s). Such additional PDP context(s) shall be established as secondary PDP contexts associated to the PDP context used for signalling.

The P-CSCF shall indicate to the UE within the ~~SIP/~~SDP according to draft-ietf-mmusic-fid-06 [48] and draft-camarillo-mmusic-separate-streams-00 [49] if a separate PDP Context is required for a media component~~ as per procedures defined in 3GPP TS 23.228 [7]~~. The UE shall establish an additional PDP context for a media component if so indicated by the received SDP~~P-CSCF~~. Media streams from different SIP sessions shall be transported in different PDP contexts.

The UE shall pass the media authorizsation token received from the P-CSCF in the 183 (Session Progress) response to an INVITE request at originating setup or in the INVITE request at terminating setup to the GGSN by inserting it within the Traffic Flow Template IE in the ACTIVATE SECONDARY PDP CONTEXT REQUEST message or the MODIFY PDP CONTEXT REQUEST message~~at PDP Context activation/modification~~. The SIP extensions for media authorization is described in RFC 3313 [29].

In order to identify to the GGSN which flow(s) (identified by m-lines within the SDP) are to be transferred within a particular PDP context, the UE shall set the flow identifier(s) within the Traffic Flow Template IE in the ACTIVATE SECONDARY PDP CONTEXT REQUEST message or the MODIFY PDP CONTEXT REQUEST message~~at PDP Context activation modification~~. Detailed description of how the flow identifiers are constructed is provided in 3GPP TS 29.207 [12].

Detailed description of how the media authorization token and flow identifiers are carried in the Traffic Flow Template IE is provided in 3GPP TS 24.008 [8].