                 SDP Superimposition Grouping framework
              draft-abhishek-mmusic-superimposition-grouping-00

Abstract

   This document defines semantics that allow for signaling a new SDP
   group "S" for superimposed media in an SDP session. The "S"
   attribute can be used by the application to relate all the
   superimposed media streams enabling them to be added as an overlay on
   top of any media stream. The superimposition grouping semantics is
   required, if the media data is separate and transported via different
   sessions.

Status of This Memo

   This Internet-Draft is submitted in full conformance with the
   provisions of BCP 78 and BCP 79.

   Internet-Drafts are working documents of the Internet Engineering
   Task Force (IETF). Note that other groups may also distribute
   working documents as Internet-Drafts. The list of current Internet-
   Drafts is at https://datatracker.ietf.org/drafts/current/.

   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."

   This Internet-Draft will expire on June 14, 2021.

Table of Contents

1. Introduction

   Media superimposition can be defined to be a visual media
   (video/image/text) which is superimposed on top of an already
   existing visual media such that the resulting foreground and
   background media can be displayed simultaneously. Superimposition
   can be recursive in that a visual media that is superimposed against
   its background can, in turn, be the background of a another
   superimposed visual media. The superimposed visual media displayed
   over a background media content may vary in transparency. Examples
   of video superimposition include real-time multi-party gaming, where
   these superimposed media maybe used to provide additional details or
   stats about each player, or multi-party teleconferencing where visual
   media from users in the teleconference may be superimposed on a
   background media or over each other. An example is shown in figure
   below, where three foreground media has been superimposed over a
   background media with one foreground media being partly superimposed
   over another foreground media.

```
    _____
   |  Background media          |
   |   _____                 |
   |  | Media A |               |
   |  |_____|               |
   |                _____ |
   |         _____|__ Media B||
   |        |Media  |__|_____||
   |        |_C_____|        |
    ----------------------------
```
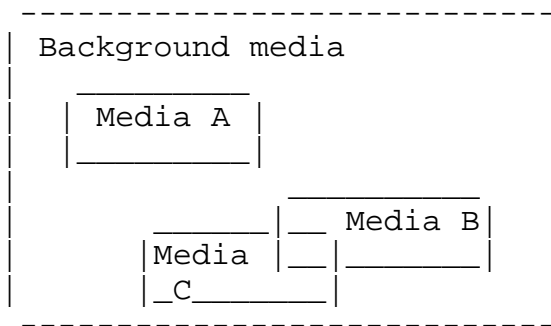
Figure 1: A example of media superimpostion

SDP is being predominantly used for describing the format for
multimedia communication sessions. Many SDP-based systems use open
standards such as RTP [RFC3550] for media transport and and SIP
[RFC3261] for session setup and control. An SDP session may contain
more than one media description with each media description
identified by "m"=line. Each line denotes a single media stream. If
multiple visual media lines are present in a session, at present,
their spatial relationship at the rendering device is undefined.
This memo introduces a mechanism in which certain rendering
information becomes available. The rendering information herein is
limited to the foreground/background relationship of each grouped
media vis-a-vis each other, and optionally a transparency value.
Where, spatially, the media is rendered is not covered by this memo
and is in many application scenarios a function of the user
interface. However, the superimposition grouping as described below
enables a compliant receiver/renderer implementation to know the
relative relevance of the visual media as coded by the sender(s) and,
in a compliant implementation, observed by the renderer through
superimposition.

When multiple superimposed streams are transmitted within a session,
the receiver needs to be able to relate the media streams to each
other. This is achieved by the SDP grouping framework [RFC5888] by
using the "group" attribute that groups different "m" lines in a
session. By using a new superimpose group semantic defined in this
memo, a group's media streams can be uniquely identified across
multiple SDP descriptions exchanged with different receivers, thereby
identifying the streams in terms of their role in the session
irrespective of its media type and transport protocol. These
superimposed streams within the group may be multiplexed based on the
guidelines defined in [draft-ietf-avtcore-multiplex-guidelines-12].

This document describes a new SDP group semantics for grouping the
superimposition in an SDP session. An SDP session description

consists of one or multiple media lines know as "m" lines which can
be identified by a token carried in a "mid" attribute. The SDP
session describes a session-level group level attribute that groups
different media lines using a defined group semantics. The semantics
defined in this memo is to be used in conjunction with "The Session
Description Protocol (SDP) Grouping Framework"[RFC5888].

2. Terminology

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 [RFC2119].

3. Superimposition Group Identification Attribute

The "superimposition media stream identification" attribute is used
to identify the relationship of superimposed media streams within a
session description. In a superimposition group, the media lines MAY
have different media formats. Its formatting in SDP [RFC4566] is
described by the following Augmented Backus-Naur Form (ABNF)
[RFC5234]:

```
mid-attribute = "a=mid:" identification-tag
identification-tag = token
                    ; token is defined in RFC4566
```

This documents defines a new group semantics "S" identification media
attribute, which is used to identify super group media streams within
a session description. It is used for grouping the foreground media
streams to be superimposed on top of a background media stream
together within a session. An application that receives a session
description that contains "m" lines grouped together using "S"
semantics MUST superimpose the corresponding media streams on top of
the background media stream. The ordering of the "m" lines is
significant: assuming the "m" lines to be counted from 0 to n, for
each k within 0..n, a reconstructed sample of the k-th media is
superimposed (while perhaps applying an alpha transparency value) on
the 0 to k-th reconstructed samples in the same spatial position.

4. Use of group and mid

All group and mid attributes MUST follow the rules defined in
[RFC5888]. The "mid" attribute MUST be used for all "m" lines
covering visual media within a session description for which a
foreground/background relationship is to be defined. The foreground/
background relationship of visual media within a session description
that is not covered in a group is undefined. No more than one group
MUST be used within one session. If the identification-tags

associated with "a=group" lines do not map to any "m" lines, it MUST
be ignored.

```
group-attribute ="a=group:" semantics
                 *(SP identification-tag)
semantics = "S" / semantics-extension
semantics-extension = token
                     ; token is defined in RFC4566
```

5. "transparency" Attribute for Superimposition Group Identification
    Attribute

   This memo defines a new media-level attribute, "transparency", with
   the following ABNF [RFC5234]. The identification-tag is defined in
   [RFC5888].

```
    transparency-attribute =
                "a=transparency:" transparency-tag
    transparency-tag =tranparency-value *("," tranparency-value) CRLF
    transparency-value= alpha
```

   Alpha describes the transparency for the foreground media stream. It
   is identified by its transparency-tag values in the transparency-
   attribute. It could be an integer with values between 0 and 100.
   This is an informative value. Details of interpretion to be left
   open to the renderer, expect that a value of 0 means foreground media
   is opaque and value of 100 means that it is transparent.

6. Example of S

   The following example shows a session description for superimposed
   media stream in an SDP session. The "group" line indicates that the
   "m" lines with tokens 1 and 2 are grouped for the purpose of
   superimpositon and intended to be overlaid on top of a background
   video.

   In the example shown below, two superimposed media streams are being
   transmitted. Both media types are video with transparency attribute
   ("transparency"). The current focus of the draft is defining a group
   semantics for superimposed media stream. The relationship between
   the background and foreground media stream maybe defined in the
   future version of the draft.

```
v=0
o=Alice 292742730 29277831 IN IP4 233.252.0.74
c=IN IP4 233.252.0.79
t=0 0
a=group:S 1 2
m=video 30000 RTP/AVP 31
a=mid:1
a= transparency: 17
m=video 30002 RTP/AVP 31
a=mid:2
a= transparency: 35
```

7. Relationship with CLUE (informative)

   Edt. Note: maybe we remove this section later once there is a general
   understanding why CLUE in its current form is unsuitable. The CLUE
   framework [I-D.ietf-clue-framework] and its associated suite of I-Ds
   and RFCs describe a telepresence framework that, at the first glance
   seems to have a lot in common with the technology proposed herein.
   CLUE defines captures (camera ports), and their geo-spatial
   relationship to each other. A render can use this information to put
   the reconstructed samples of the streams from the various captures
   into a suitable arrangement such that visually pleasant rendering can
   be achieved. However, CLUE does not describe the relative relevance
   of the captures. For that reason, CLUE would need to be extended in
   a spirit very similar to the one described in this memo to achieve
   the desired functionality. CLUE has not seen wide deployment outside
   its intended key application (large room, multiple camera
   telepresence systems). It's not reasonable to assume that small
   systems would willingly implement the overhead the (comparatively
   complex) CLUE protocols require when a simple SDP extension can serve
   the same purpose.

8. Security Considerations

   All security considerations as defined in [RFC5888] apply:

   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 destination of its
   choosing.

   Integrity mechanisms provided by protocols used to exchange session
   descriptions and media encryption can be used to prevent this attack.
   In SIP, Secure/Multipurpose Internet Mail Extensions (S/MIME)
   [RFC8550] and Transport Layer Security (TLS) [RFC8446] can be used to

protect session description exchanges in an end-to-end and a hop-
byhop fashion, respectively.

9. IANA Considerations

The following contact information shall be used for all registrations
included here:

Contact: Rohit Abhishek
                  email: rabhishek@rabhishek.com
                  tel : +1-816-585-7500

This document defines a new SDP group semantics for media
superimposition for a SDP session. This attribute can be used by the
application to group all the foreground media to be superimposed on a
background media in a session. Semantics values to be used with this
framework should be registered by the IANA following the Standards
Action policy [RFC8126]. This document adds a new group semantics
and follows the registry group defined in [RFC5888].

The following semantics needs to be registered by IANA in Semantics
for the "group" SDP Attribute under SDP Parameters.

```
Semantics                 Token            Reference
-------------------------------------------------
Superimposition             S                RFCXXXX
```

The "S" attribute is used to group different foreground media streams
to be superimposed on a background media stream . Its format is
defined in Section 3.

The SDP media-level attribute "transparency" needs to be registered
by IANA Semantics for "att-field (media-level only)". The
registration procedure in [RFC4566] applies.

SDP Attribute ("att-field (media level only)"):

                  Attribute name: transparency
                  Long form: superimposition transparency
                  Type of name: att-field
                  Type of attribute: media level only
                  Subject to charset: no
                  Purpose: RFC 5583
                  Reference: RFC 5583
                  Values: alpha

The IANA Considerations section of the RFC MUST include the following
information, which appears in the IANA registry along with the RFC
number of the publication.

o A brief description of the semantics.

o Token to be used within the "group" attribute. This token may be
   of any length, but SHOULD be no more than four characters long.

o Reference to a standards track RFC.

10. References

10.1. Normative References

   [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate
              Requirement Levels", BCP 14, RFC 2119,
              DOI 10.17487/RFC2119, March 1997,
              <https://www.rfc-editor.org/info/rfc2119>.

   [RFC3261] Rosenberg, J., Schulzrinne, H., Camarillo, G., Johnston,
              A., Peterson, J., Sparks, R., Handley, M., and E.
              Schooler, "SIP: Session Initiation Protocol", RFC 3261,
              DOI 10.17487/RFC3261, June 2002,
              <https://www.rfc-editor.org/info/rfc3261>.

   [RFC3550] Schulzrinne, H., Casner, S., Frederick, R., and V.
              Jacobson, "RTP: A Transport Protocol for Real-Time
              Applications", STD 64, RFC 3550, DOI 10.17487/RFC3550,
              July 2003, <https://www.rfc-editor.org/info/rfc3550>.

   [RFC4566] Handley, M., Jacobson, V., and C. Perkins, "SDP: Session
              Description Protocol", RFC 4566, DOI 10.17487/RFC4566,
              July 2006, <https://www.rfc-editor.org/info/rfc4566>.

   [RFC5234] Crocker, D., Ed. and P. Overell, "Augmented BNF for Syntax
              Specifications: ABNF", STD 68, RFC 5234,
              DOI 10.17487/RFC5234, January 2008,
              <https://www.rfc-editor.org/info/rfc5234>.

   [RFC5888] Camarillo, G. and H. Schulzrinne, "The Session Description
              Protocol (SDP) Grouping Framework", RFC 5888,
              DOI 10.17487/RFC5888, June 2010,
              <https://www.rfc-editor.org/info/rfc5888>.

   [RFC8126] Cotton, M., Leiba, B., and T. Narten, "Guidelines for
              Writing an IANA Considerations Section in RFCs", BCP 26,
              RFC 8126, DOI 10.17487/RFC8126, June 2017,
              <https://www.rfc-editor.org/info/rfc8126>.

   [RFC8446] Rescorla, E., "The Transport Layer Security (TLS) Protocol
              Version 1.3", RFC 8446, DOI 10.17487/RFC8446, August 2018,
              <https://www.rfc-editor.org/info/rfc8446>.

   [RFC8550] Schaad, J., Ramsdell, B., and S. Turner, "Secure/
              Multipurpose Internet Mail Extensions (S/MIME) Version 4.0
              Certificate Handling", RFC 8550, DOI 10.17487/RFC8550,
              April 2019, <https://www.rfc-editor.org/info/rfc8550>.

10.2. Informative References

   [draft-ietf-avtcore-multiplex-guidelines-12]
              Westerlund, M., Burman, B., Perkins, C., Alvestrand, H.,
              and R. Even, "Guidelines for using the Multiplexing
              Features of RTP to Support Multiple Media Streams", draft-
              ietf-avtcore-multiplex-guidelines-12 (work in progress),
              June 2020.

   [I-D.ietf-clue-framework]
              Duckworth, M., Pepperell, A., and S. Wenger, "Framework
              for Telepresence Multi-Streams", draft-ietf-clue-
              framework-25 (work in progress), January 2016.

Authors' Addresses

   Rohit Abhishek
   Tencent
   2747 Park Blvd
   Palo Alto 94588
   USA

   Email: rabhishek@rabhishek.com

   Stephan Wenger
   Tencent
   2747 Park Blvd
   Palo Alto 94588
   USA

   Email: stewe@stewe.org