MMUSIC Working Group   J. Hautakorpi
Internet-Draft   G. Camarillo
Expires: August 15, 2005   Ericsson
                                                      February 14, 2005

The SDP (Session Description Protocol) Content Attribute
draft-hautakorpi-mmusic-sdp-media-content-00.txt

Status of this Memo

Copyright Notice

Abstract

   This document defines a new Session Description Protocol (SDP)
   media-level attribute, 'content'. The 'content' attribute defines
   the content of the media stream in more detailed level than the media
   description line. The sender of an SDP session description can
   attach the 'content' attribute to one or more media streams. The
   receiving application can then treat each media stream differently
   (e.g., show it on a big screen or small screen) based on their

   content.

Table of Contents

1. Introduction

   SDP [1] is a protocol that is intended for describing multimedia
   sessions for the purposes of session announcement, session
   invitation, and other forms of multimedia session initiation. One of
   the most typical use cases of SDP is the one where it is used with
   SIP [3].

   There are situations where one application receives several similar
   media stream which are described in an SDP session description. The
   media streams can be similar in a sense that their content cannot be
   distinguished from each other just by examining the media description
   lines (e.g., two video streams). The 'content' attribute is needed,
   so that the receiving application can treat each media stream
   appropriately based on its content.

   This specification defines the SDP 'content' media-level attribute,
   which provides more information about the media stream than the 'm'
   line in an SDP session description.

2. Motivation for the New Content Attribute

   Currently, SDP does not provide any means to describe what is the
   content of a media stream (e.g., speaker's image, slides, sign
   language) in a form that the application can understand. Of course
   the end user can see the content of the media stream, and read its
   title, but the application cannot understand what the media stream
   contains.

   The application that is receiving multiple similar (e.g., same type
   and format) media stream needs, in some cases, to know what is the
   content of those streams. This kind of situation occurs for example
   in cases where presentation slides, speakers image and sign language
   are transported as separate media streams. It would be desirable
   that the receiving application could distinguish them in a way that
   it could handle them automatically in appropriate manner.

```
            +-------------------------------------+
            |+-----------++---------------------+|
            ||          ||                       ||
            || speaker's ||                       ||
            ||   image    ||                       ||
            ||          ||                       ||
            |+-----------+|    presentation       ||
            |+-----------+|       slides          ||
            ||          ||                       ||
            ||   sign    ||                       ||
            ||  language ||                       ||
            ||          ||                       ||
            |+-----------++---------------------+|
            +-------------------------------------+
```
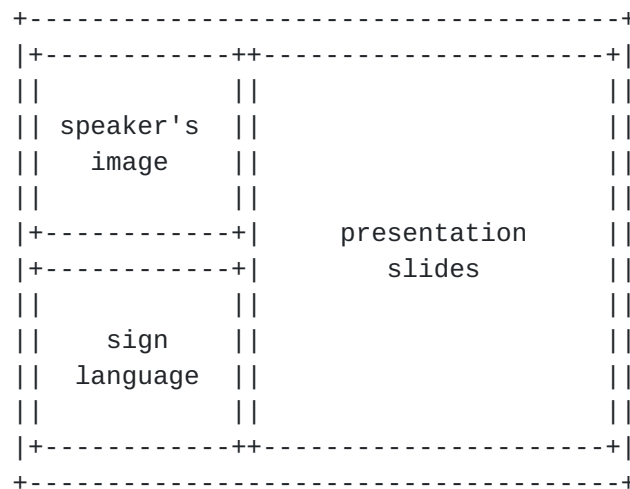
                    Figure 1: Application's screen

   The Figure 1 presents a screen of a typical communication
   application. The 'content' attribute enables the application to make
   the decision on where to show each media stream. From end user's
   perspective, it is desirable that the user does not need to arrange
   media stream every time the media session starts.

   The 'content' attribute could also be used in more complex
   situations. This kind of complex situation could be e.g., an
   application that is controlling the equipments in the auditorium.
   Auditorium can have many different output channels for the video
   (main screen and two smaller screens) and the audio (main speakers,
   headsets for the participants). In this kind of environment, a lot
   of interaction from the end user who is operating the application
   would be required. So, the possibility for the application to handle
   the media stream without end users input is greatly emphasized.

3. The Content Attribute

   This specification defines a new media-level value attribute,
   'content'. Its formatting in SDP is described by the following BNF
   [2]:

```
     content-attribute   = "a=content:" mediacnt
     mediacnt            = "slides" / "speaker" / "sl" / "main-cam"
                           / "alt-cam" / "main-audio" / "alt-audio"
                           / mediacnt-ext
     mediacnt-ext        = token
```

   The 'content' attribute contains a token, which is attached to a
   media stream by a sending application. It describes the content of

the trasmitted media stream to the receiving application.

There are six different pre-defined values for the 'content'
attribute. Also other values can be defined in the future. The
pre-defined values are:

slides: This is a media stream that includes presentation slides.
   The media type can be e.g., a video stream or a set of instant
   message with pictures. Typical use case for this is e.g., online
   seminars and courses.
speaker: This is a image from the speaker. The media can be e.g., a
   video stream or a still image. Typical use case for this is e.g,
   online seminars and courses.
sl: This means that the media stream contains sign language. The
   media type is a videa stream. Typical use case for this is the
   one where the audio stream is translated into sign language.
main-cam: This means that the video stream is taken from the main
   camera. Typical use case for this is e.g., a concert, where the
   camera is shooting the performer. This is similar to
   'presentation' role in H.239 [5].
alt-cam: This means that the video stream is taken from the
   alternative camera. Typical use case for this is e.g., a concert,
   where the camera is shooting the crowd. This is similar to 'live'
   role in H.239.
main-audio: This is the main audio stream. Typically this can be
   used in situations, where there is a separate ambient sound and
   the main sound. The main sound could be e.g., a sound of a rare
   bird.
alt-audio: This is the alternative audio stream. Typically this can
   be used in situations, where there is a separate ambient sound and
   the main sound. The alternative sound could be e.g., a sound of a
   jungle.

The media used with the first two pre-defined values 'slides' and
'speaker' can be in many different kinds of formats. The following
three pre-defined values 'sl', 'main-cam' and 'alt-cam' can be used
only on the context of video streams. The last two pre-defined
values 'main-audio' and 'alt-audio' can be used only with audio
stream.

Open issue: Can multiple 'content' attributes, or attribute values,
be defines for a single media stream?

4. The Content Attribute in the Offer/Answer Model

This specification does not define a means to discover whether or not
the peer endpoint understands the 'content' attribute, because
'content' values are informative only at the offer/answer model [4]

   level. The fact that the peer endpoint does not understands the
   'content' attribute, does not keep the media session from being
   established. The only consequence is that interaction from the
   receiving end user may be required.

[5](#). **Example**

   The following is an example of the SDP session description that uses
   the 'content' attribute:

      v=0
      o=Alice 292742730 29277831 IN IP4 131.163.72.4
      s=Second lecture from information technology
      c=IN IP4 131.164.74.2
      t=0 0
      m=video 52886 RTP/AVP 31
      a=rtpmap:31 H261/9000
      a=content:slides
      m=audio 53334 RTP/AVP 31
      a=rtpmap:31 H261/9000
      a=content:speaker
      m=audio 54132 RTP/AVP 31
      a=rtpmap:31 H261/9000
      a=content:sl

[6](#). **Security Considerations**

   TBD.

[7](#). **IANA Considerations**

   This document defines a new 'content' attribute for SDP. It also
   defines seven possible values for it.

   Contact name: Jani Hautakorpi Jani.Hautakorpi@ericsson.com.

   Attribute name: 'content'.

   Type of attribute   Media level.

   Subject to charset: Not.

   Purpose of attribute: The 'content' attribute gives information from
      the content of the media stream to the receiving application.

    Allowed attribure values: "slides", "speaker", "sl", "main-cam",
                              "alt-cam", "main-audio", "alt-audio",
                              and a token.

    Entries to the registry:

    Value of 'content' attribute   Reference   Description
    ----------------------------   ---------   -----------
    slides   RFC xxxx   Presentation slides
    speaker   RFC xxxx   Image from the speaker
    sl   RFC xxxx   Sign language
    main-cam   RFC xxxx   Main media stream
    alt-cam   RFC xxxx   Alternative media stream
    main-audio   RFC xxxx   Main audio stream
    alt-audio   RFC xxxx   Alternative audio stream

    Open issue: Which policy are we going to use for defining new
    extension tokens?

8. Acknowledges

    Aroud van Wijk and Roni Even provided valuable ideas for this
    document.

9. References

9.1 Normative References

    [1]   Handley, M., "SDP: Session Description Protocol",
          draft-ietf-mmusic-sdp-new-23 (work in progress), December 2004.

    [2]   Crocker, D. and P. Overell, "Augmented BNF for Syntax
          Specifications: ABNF", RFC 2234, November 1997.

9.2 Informational References

    [3]   Rosenberg, J., Schulzrinne, H., Camarillo, G., Johnston, A.,
          Peterson, J., Sparks, R., Handley, M. and E. Schooler, "SIP:
          Session Initiation Protocol", RFC 3261, June 2002.

    [4]   Rosenberg, J. and H. Schulzrinne, "An Offer/Answer Model with
          Session Description Protocol (SDP)", RFC 3264, June 2002.

    [5]   ITU-T, "Infrastructure of audiovisual services - Systems
          aspects; Role management and additional media channels for
          H.300-series terminals", Series H H.239, July 2003.

Authors' Addresses

    Jani Hautakorpi
    Ericsson
    Hirsalantie 11
    Jorvas 02420
    Finland

    EMail: Jani.Hautakorpi@ericsson.com

    Gonzalo Camarillo
    Ericsson
    Hirsalantie 11
    Jorvas 02420
    Finland

    EMail: Gonzalo.Camarillo@ericsson.com

Intellectual Property Statement

   The IETF takes no position regarding the validity or scope of any
   Intellectual Property Rights or other rights that might be claimed to
   pertain to the implementation or use of the technology described in
   this document or the extent to which any license under such rights
   might or might not be available; nor does it represent that it has
   made any independent effort to identify any such rights. Information
   on the procedures with respect to rights in RFC documents can be
   found in BCP 78 and BCP 79.

   Copies of IPR disclosures made to the IETF Secretariat and any
   assurances of licenses to be made available, or the result of an
   attempt made to obtain a general license or permission for the use of
   such proprietary rights by implementers or users of this
   specification can be obtained from the IETF on-line IPR repository at
   http://www.ietf.org/ipr.

   The IETF invites any interested party to bring to its attention any
   copyrights, patents or patent applications, or other proprietary
   rights that may cover technology that may be required to implement
   this standard. Please address the information to the IETF at
   ietf-ipr@ietf.org.

Disclaimer of Validity

   This document and the information contained herein are provided on an
   "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS
   OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY AND THE INTERNET
   ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED,
   INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE
   INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED
   WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Copyright Statement

   Copyright (C) The Internet Society (2005). This document is subject
   to the rights, licenses and restrictions contained in BCP 78, and
   except as set forth therein, the authors retain all their rights.

Acknowledgment

   Funding for the RFC Editor function is currently provided by the
   Internet Society.