The SDP (Session Description Protocol) Content Attribute

draft-ietf-mmusic-sdp-media-content-06.txt

Status of this Memo

By submitting this Internet-Draft, each author represents that any applicable patent or other IPR claims of which he or she is aware have been or will be disclosed, and any of which he or she becomes aware will be disclosed, in accordance with Section 6 of BCP 79.

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

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

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

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

This Internet-Draft will expire on March 26, 2007.

Copyright Notice

Copyright (C) The Internet Society (2006).

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 its content.
Table of Contents

1. Introduction ................................................. 3
2. Terminology ................................................. 3
3. Related Techniques ........................................... 3
4. Motivation for the New Content Attribute .................... 4
5. The Content Attribute ........................................ 5
6. The Content Attribute in the Offer/Answer Model ............. 6
7. Examples ................................................... 7
8. Operation with SMIL ......................................... 8
9. Security Considerations ...................................... 8
10. IANA Considerations ........................................ 8
11. Acknowledgements .......................................... 9
12. References ................................................ 10
   12.1. Normative References ................................. 10
   12.2. Informational References ............................. 10
Authors’ Addresses ............................................. 11
Intellectual Property and Copyright Statements ................ 12
1. Introduction

The Session Description Protocol (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 the Session Initiation Protocol (SIP) [5].

There are situations where one application receives several similar media streams which are described in an SDP session description. The media streams can be similar in the sense that their content cannot be distinguished just by examining their 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.

The main purpose of this specification is to allow applications to take automated actions based on the ‘content’ attributes. However, this specification does not define those actions. Consequently, two implementations can behave completely differently when receiving the same ‘content’ attribute.

2. Terminology

In this document, the key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" are to be interpreted as described in BCP 14, RFC 2119 [3] and indicate requirement levels for compliant implementations.

3. Related Techniques

The ‘label’ attribute [10] enables a sender to attach a pointer to a particular media stream. The name space of the ‘label’ attribute itself is unrestricted; so, in principle it could also be used to convey information about the content of a media stream. However, in practice, this is not possible because of the need for backward compatibility. Existing implementations of the ‘label’ attribute already use values from that unrestricted namespace in an application-specific way. So, it is not possible to reserve portions of the ‘label’ attribute’s namespace without possible conflict with
already-used application-specific labels.

It is possible to assign semantics to a media stream with an external document that uses the ‘label’ attribute as a pointer. The downside of this approach is that it requires an external document. Therefore, this kind of mechanism is only applicable to special use cases where such external documents are used (e.g., centralized conferencing).

Yet another way to attach semantics to a media stream is to use the ‘i’ SDP attribute, defined in [1]. However, values of the ‘i’ attribute are intended for human users and not for automata.

4. 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, the speaker’s 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 an 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 makes it possible for the application to decide where to show each media stream. From an end user’s perspective, it is desirable that the user does not need to arrange media stream every time a new media session starts.

The ‘content’ attribute could also be used in more complex situations. An example of such a complex situation is an application controlling equipment in an auditorium. An auditorium can have many different output channels for video (e.g., main screen and two smaller screens) and audio (e.g., main speakers, headsets for the participants). In this kind of environment, a lot of interaction from the end user who operates the application would be required in absence of cues from a controlling application. The ‘content’ attribute would make it possible, for example, for an end user needs to specify, only once, which output each media stream of a given session should use. The application could automatically apply the same media layout for subsequent sessions. So, the ‘content’ attribute can help to reduce the amount of required end user interaction considerably.

5. 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-tag
mediacnt-tag = mediacnt *("," mediacnt)
mediacnt = "slides" / "speaker" / "sl" / "main"
         / "alt" / mediacnt-ext
mediacnt-ext = token
```

The ‘content’ attribute contains a token, which MAY be attached to a media stream by a sending application. An application MAY attach a content attribute to any media stream it describes. That attribute contains one or more tokens describing the content of the transmitted media stream to the receiving application.

This document provides a set of pre-defined values for the ‘content’ attribute. Other values can be defined in the future. The pre-defined values are:
slides: the media stream includes presentation slides. The media type can be, for example, a video stream or a number of instant messages with pictures. Typical use cases for this are online seminars and courses. This is similar to the 'presentation' role in H.239 [12].

speaker: the media stream contains the image of the speaker. The media can be, for example, a video stream or a still image. Typical use case for this are online seminars and courses.

sl: the media stream contains sign language. A typical use case for this is an audio stream that is translated into sign language, which is sent over a video stream.

main: the media stream is taken from the main source. A typical use case for this is a concert where the camera is shooting the performer.

alt: the media stream is taken from the alternative source. A typical use case for this is an event where the ambient sound is separated from the main sound. The alternative audio stream could be, for example, the sound of a jungle. Another example is the video of a conference room while the main stream carries the video of the speaker. This is similar to the 'live' role in H.239.

All these values can be used with any media type. The application can make decisions on how to handle a single media stream based on both the media type and the value of the 'content' attribute. Therefore the situation where one value of 'content' attribute occurs more than once in a single session descriptor is not problematic.

6. 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 just informative at the offer/answer model [8] level. The fact that the peer endpoint does not understand the 'content' attribute does not keep the media session from being established. The only consequence is that end user interaction on the receiving side may be required to direct the individual media streams appropriately.

Since the 'content' attribute does not have to be understood, an SDP answer MAY contain 'content' attributes even if none were present in the offer. Similarly, the answer MAY contain no 'content' attributes even if they were present in the offer. Furthermore, the values of 'content' attributes does not need to match in an offer and an
answer.

The ‘content’ attribute can also be used in scenarios where SDP is used in a declarative style. For example, ‘content’ attributes can be used in SDP session descriptors that are distributed with Session Announcement Protocol (SAP) [9].

7. Examples

There are two examples in this section. The first example, shown below, uses a single ‘content’ attribute value per media stream:

```
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=video 53334 RTP/AVP 31
a=rtpmap:31 H261/9000
a=content:speaker
m=video 54132 RTP/AVP 31
a=rtpmap:31 H261/9000
a=content:sl
```

The second example, below, shows a case where there is more than one ‘content’ attribute value per media stream. The difference with the previous example is that now the conferencing system might automatically mix the video streams from the presenter and slides:

```
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,speaker
m=video 54132 RTP/AVP 31
a=rtpmap:31 H261/9000
a=content:sl
8. Operation with SMIL

The values of 'content' attribute, defined in Section 5, can also be used with SMIL [11]. SMIL contains a 'param' element, which is used for describing the content of a media flow. However, this 'param' element, like 'content' attribute, provides application specific description of media content.

Details on how to use the values of the 'content' attribute with SMIL’s 'param' element are outside the scope of this specification.

9. Security Considerations

An attacker may attempt to add, modify, or remove 'content' attributes from a session description. Depending on how an implementation chooses to react to the presence or absence of a given 'content' attribute, this could result in an application behaving in an undesirable way. So, it is strongly RECOMMENDED that integrity protection be applied to the SDP session descriptions.

Integrity protection can be provided for session description carried in SIP [5] e.g., by using S/MIME [6] or Transport Layer Security (TLS) [7].

It is assumed that values of 'content' attribute do not contain data that would be truly harmful if it is exposed to an possible attacker. It must be noted that the initial set of values does not contain any data that would require confidentiality protection. However, S/MIME and TLS can be used to protect confidentiality, if needed.

10. IANA Considerations

This document defines a new 'content' attribute for SDP. It also defines an initial set of values for it. Some general information regarding 'content' attribute is presented in the following:
Contact name: Jani Hautakorpi Jani.Hautakorpi@ericsson.com.

Attribute name: 'content'.

Type of attribute Media level.

Subject to charset: No.

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

Allowed attribute values: "slides", "speaker", "sl", "main", "alt", and any other registered values.

The IANA is requested to create a subregistry for ‘content’ attribute values under the Session Description Protocol (SDP) Parameters registry. The initial values for the subregistry are presented in the following, and IANA is requested to add them into its database:

<table>
<thead>
<tr>
<th>Value of 'content' attribute</th>
<th>Reference</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>slides</td>
<td>RFC xxxx</td>
<td>Presentation slides</td>
</tr>
<tr>
<td>speaker</td>
<td>RFC xxxx</td>
<td>Image from the speaker</td>
</tr>
<tr>
<td>sl</td>
<td>RFC xxxx</td>
<td>Sign language</td>
</tr>
<tr>
<td>main</td>
<td>RFC xxxx</td>
<td>Main media stream</td>
</tr>
<tr>
<td>alt</td>
<td>RFC xxxx</td>
<td>Alternative media stream</td>
</tr>
</tbody>
</table>

Note for the RFC Editor: ‘RFC xxxx’ above should be replaced by a reference to the coming RFC number of this draft.

As per the terminology in RFC 2434 [4], the registration policy for new values for the ‘content’ parameter shall be ‘Specification Required’.

If new values for ‘content’ attribute are specified in the future, they should consist of a meta description of the contents of a media stream. New values for ‘content’ attribute should not describe things like what to do in order to handle a stream.

11. Acknowledgements

Authors would like to thank Arnoud van Wijk and Roni Even, who provided valuable ideas for this document. We wish to thank also Tom Taylor for a thorough review.
12. References

12.1. Normative References


12.2. Informational References


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