

SDP：会话描述协议

摘要

本规范定义了会话描述协议（SDP）。SDP 是为了描述多媒体会话而设计的，主要用来描述会话通告、会话邀请或者其他形式的多媒体会话初始化。

1 引言

当初始化多媒体会议、VoIP 电话，视频流或者其他会话的时候，需要对参与者提供媒体信息、传输地址以及其他与会话相关的描述性的元数据。

SDP为描述这些信息提供了一个标准的形式，并且与这些信息是以什么方式传送的无关。SDP只是一个单纯的会话描述的格式，SDP并不包含传输协议。在不同的场合，SDP可以应用于不同的传输协议，如SAP、SIP、RTP、使用MIME扩展的电子邮件和HTTP等。

SDP被设计成通用的，因此它可以广泛用于有关网络的各种应用。然而，它并不能用于协商会话内容或者媒体编码，这超出了SDP的功能范围。

这个规范废除了RFC2327和RFC3266，在第10部分，列出了这个规范和RFC2327和RFC3266的不同。

2 术语

下面是在本文档中会用到的术语，并且在本文档的上下文中会有相应的特殊含义。

会议：一个多媒体会议是一个集合，这个集合包含两个或者更多的通信用户及其用于通信的软件。

会话：一个多媒体会话是一个集合，这个集合包含发送者和接受者，以及他们之间所通信的媒体流。多媒体会议是多媒体会话的一个例子。

会话描述：它的格式是well-defined格式，能够传达足够的信息以用于发现和参与一个多媒体会话。

诸如"MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", 和 "OPTIONAL" 这些关键字符合RFC2119中的规则。

3 SDP 应用的例子

3.1 会话初始化

SIP是一个应用层的控制协议，可以用于创建、修改和终止会话，包含互联网多媒体会议、互联网电话和多媒体分发。用于创建会话的**SIP**消息包含**SDP**，使得参与者可以协商双方兼容的媒体类型。当用于**SIP**时，使用**SDP**的offer/answer提供了一个有限的框架。

3.2 流媒体

RTSP是一个应用层的协议，用来控制实时数据的传输的。**RTSP**提供了一个可扩展的框架用来控制和按要求传输例如音频和视频的实时数据。**RTSP**客户端和服务端协商一系列的参数用于媒体传输，其中一部分参数是用**SDP**进行描述的。

3.3 邮件和万维网

传输**SDP**的其他的方法包括**email**和万维网。**email**和万维网分发的媒体类型都是"**application/sdp**"。这可以使万维网的客户端和邮件接收方以标准的方式自动登陆应用。

注意，接收方不一定能收到只使用**email**或者万维网的多播会话的通告。主要是因为多播会话可能限制于某一个范围内，而接收方则可能在这个范围之外。

3.4 多播会话通告

为了帮助多播多媒体会议或者其他的多播会话做宣传，以及把会话相关的**setup**信息传达给参与者，有可能会采用分发会话目录的形式。这种方式的一个例子是发送一个包含会话描述信息的数据包给一个众所周知的多播组。其他会话目录可以收到这些信息，因此远端的参与者可以参与会话。

用于实现这种分发目录的其中一个协议是**SAP**。而**SDP**则为这些会话通告提供了一种会话描述格式。

4 要求和建议

SDP的目的是传达多媒体会话中有关媒体流的信息，使得会话描述信息的接收方可以参与会话。尽管**SDP**被设计成通用的，可以用于各种网络环境，但它通常用于互联网服务。媒体流是可以多对多的，会话也不要求持续活跃。

到目前为止，互联网上基于多播的电话会议与其他形式的电话会议有很大不同，原因在于互联网上基于多播的电话会议的接收者只要接收到这些信息的都可以加入会话，除非这些信息是通过加密传输的。在这种环境下，**SDP**主要有两个作用。一个就是保活，另外一

个就是提供足够的信息加入会话。在单播环境下，只有用到后面那个作用。

一个SDP会话描述包含以下内容：

- 会话名字和目的
- 会话持续的时间
- 会话的媒体类型
- 接受这些媒体数据所需要的信息，如地址、端口、格式等

由于参与会话的资源可能是受限的，因此，一些额外的信息也是比较期望的：

- 会话所占用的带宽
- 会话拥有者的联系信息

一般来说，SDP 必须传达足够的信息使得应用可以加入会话，以及向非参与者宣告所使用的资源。

4.1 媒体和传输信息

一个SDP会话描述包含以下媒体信息：

- 媒体类型：视频或者音频
- 传输协议：UDP或者H.320
- 媒体格式信息：H.263或者MPEG4

除了媒体格式和传输协议之外，SDP传达了地址和端口信息。对于一个IP多播会话，包含如下信息：

- 媒体的多播组地址
- 媒体的传输端口

这个地址和端口是多播媒体流目的地址和目的端口，发送的或者接收的或者两者都有。

对于单播IP会话，包含如下信息：

- ✓ 媒体的远端IP地址
- ✓ 媒体的远端传输端口

这个地址和端口的语义取决于媒体类型和传输协议。缺省情况下，它是媒体流的源IP地址和端口。某些媒体类型会重定义这个行为，但是这是不推荐使用的。

4.2 时间信息

会话持续的时间可以是有限的，也可以是无限的。无论如何，会话会在某个特定的时间被激活。SDP可能会包含如下信息：

- 会话的一个时间列表，包含开始时间和结束时间
- 对于每一段时间，可以指定重复发生，例如每星期三的早上10点开始，持续一个小时。

这个时间信息是全球统一的，不考虑本地时域或者夏令时。

4.3 私有会话

在建立会话的时候可以建立公开的或者是私有的会话。SDP本身并不区分这两点，只是在分发私有会话的会话描述信息的时候是采用加密的形式分发的。至于加密的方法则取决于

传达SDP的机制，现有的机制如SAP和SIP，将来会有其他的形式。

如果一个会话通告是私有的，则可能是用来传输加密密钥以用于对会议的媒体信息进行解码，而且还包含足够的信息用于区分哪些媒体流使用哪种加密机制。

4.4 获取有关会话的额外的信息

一个会话描述应该包含足够的信息用于决定是否参与一个会话。SDP 可能还会以 URI 的形式指定一些额外的信息。

4.5 分类

当多个会话描述通过SAP或者其他机制进行分发的时候，有必要对有用和无用的SDP进行区分。SDP提供一种分类机制，通过属性“a=cat:”来自动完成分类功能。

4.6 国际化

SDP规范建议使用UTF-8中的ISO 10646字体集合，以便可以用于表示各种语言。但是，为了能够更好进行表述，SDP也支持其他字体集合如ISO 8859-1。这种国际化只会应用于一些文本区域，而不是整个SDP。

5 SDP 规范

一个 SDP 会话描述是通过"application/sdp"来标识的。

一个SDP会话描述全部都是采用UTF-8编码的ISO 10646字体集合的文本。SDP的域名和属性名是采用UTF-8的US-ASCII子集，而其他形式则是采用整个UTF-8编码的ISO 10646字体集合。采用文本形式主要是为了加强可移植性，使得可以应用于更多的传输形式，并且也可以让更多的机遇文本的工具箱能够生成和处理SDP。但是由于在某些场合，SDP的长度可能受限，因此编码格式会设计成紧凑的。而且，既然通告可能在一些不稳定的环境中进行传输，因此也会设计成有严格的顺序和有一定的格式规则，以便有错误产生时能够容易发现。这同样也允许在加密的传输中。如果接收端没有正确的密钥，则可以快速丢弃。

一个SDP会话描述包含若干行如下形式的文本：

<type>=<value>

其中<type>是一个区分大小写的字母的形式，<value>是一个结构化的文本，其形式取决于<type>。通常，<value>由若干分隔符隔开的字段组成，或者是一个字符串。其中"="两遍都不能有空格。

一个SDP包含一个会话层部分，后续跟着0个或者多个媒体层部分。会话层起始于“v=”行，直到第一个媒体层部分。每个媒体层都起始于“m=”行，直到下一个媒体层的开始或者直到SDP结束。通常，会话层的属性是全局的，除非被媒体层重写。

某些行是必选的，某些是可选的，但是出现的顺序必须如下所示。可选项用“*”标志。

会话描述

v=(协议版本)

o=(发起者和会话标识符)
s=(会话名字)
i=(会话信息)
u=(URI标识符)
e=(邮件地址)
p=(电话号码)
c=(连接信息，只有在所有的媒体层都有这个字段的时候，会话层属性才可以没有)
有)

b=(0个或者多个带宽信息)
一个或者多个时间描述("t="和"r="行)
z=(时域调整)
k=(加密密钥)
a=(0个或者会话属性)
0个或者多个媒体描述

时间信息

t=(会话有效的时间)
r=(0个或者多个重复时间)

媒体描述，如果存在的情况下

m=(媒体名字和传输地址)
i=(媒体标题)
c=(连接信息，如果在会话层中出现，则这里可以不出现)
b=(0个或者多个带宽信息)
k=(加密密钥)
a=(0个或者多个媒体属性)

<type>的字母为小写字母并且是不可扩展的，SDP解析器如果碰到不认识的<type>，必须忽略它。属性机制("a=")是SDP扩展的主要手段。某些属性是已经被定义好的，而另外的则是在特定的应用、媒体格式规范中才进行定义。SDP解析器必须忽略任何它不认识的属性。

一个SDP会话描述可能包含URI以指定外部的信息，主要在"u=","k="和"a="这三个字段中有所体现。这些URI在某些场合下是非关联的，避免会话描述自包含。

会话层的连接信息("c=")和属性("a=")都是全局的，除非在媒体层被重写。例如如下例子，每个媒体层都有一个"recvonly"属性。

```
v=0
o=jdoe 2890844526 2890842807 IN IP4 10.47.16.5
s=SDP Seminar
i=A Seminar on the session description protocol
u=http://www.example.com/seminars/sdp.pdf
e=j.doe@example.com (Jane Doe)
c=IN IP4 224.2.17.12/127
t=2873397496 2873404696
a=recvonly
m=audio 49170 RTP/AVP 0
m=video 51372 RTP/AVP 99
a=rtpmap:99 h263-1998/90000
```

文本字段如会话名和信息是字符串形式，其中不包含字符0x00，0x0a，0x0d。0x0d0a连起来作为一行的终止符，当然，解析器可能允许单一的字符作为终止符。如果"a=charset"

不存在，这些字符串的格式为UTF-8编码的ISO 10545字体集。

SDP可能在“o=”，“u=”，“e=”，“c=”和“a=”等字段包含域名。域名规则必须符合【1】和【2】文献的要求。IDNs必须符合ACE规则，而直接用UTF-8或者其他编码格式。

5.1 协议版本（“v=”）

v=0

“v=”字段给出了SDP的版本，当前规范的版本是0，这个版本没有小版本号。

5.2 源（“o=”）

o=<用户名> <会话ID> <会话版本> <网络类型> <地址类型> <单播地址>

“o=”字段给出了当前会话的发起者，会话标志和版本号的信息。

<用户名>是登录源主机的用户的名字,如果不能提供,则用“-”表示，用户名不能包含空格。

<会话ID>是一个数字字符串。<用户名>、<会话ID>、<网络类型>、和<单播地址>这个组合形成了表示该会话的唯一标识。<会话ID>通常用NTP格式的时戳表示。

<会话版本>是当前会话描述的版本号。如果当前会话的数据被修改时，这个版本号会递增。同样，推荐使用NTP格式的时戳。

<网络类型>是一个文本字符串，标志当前的网络类型。最初为“IN”表示“Internet”，未来可能会定义其他值。

<地址类型>是一个文本字符串，标志当前地址的类型。最初只有“IP4”和“IP6”有定义，未来可能定义其他值。

<单播地址>是创建会话的主机的地址。可以是域名，也可以是IPv4或者IPv6的IP地址。域名是首选，除非域名不可获得。如果使用IP地址，则需要使用全球唯一的IP地址，不能使用本地IP地址。

一般来说，“o=”唯一标识了当前会话描述的版本，出去子字段版本，则唯一标识了当前会话。

有时候处于某种原因，用户名和IP地址可能不想明确表示。这时候，只要能保证“o=”字段全球唯一，可以采用随机的用户名和私有的IP地址。

5.3 会话名字（“s=”）

s=<会话名字>

字段“s=”是文本类型的会话名字，每个会话描述有且仅有一个“s=”字段。“s=”字段不能为空，并且应该为ISO 10646字体。如果一个会话没有有意义的名字，则此字段的值应为一个空格。

5.4 会话信息（“i=”）

i=<会话描述>

“i=”字段提供会话的文本信息。在会话层和媒体层最多只能出现一次。如果有

“a=charset”字段，它标志了“i=”字段使用的字体集合。如果没有“a=charset”字段，则必须使用UTF-8编码的ISO 10646字体集合。

“i=”字段也可以用于每个媒体层，这是用来标志每个媒体轨的。特别在两个媒体轨的媒体类型都是相同的情况下，这个字段就显得特别有意义了。

这个字段主要是用来方便人来阅读的，并不适合自动解析。

5.5 统一资源描述符（“u=”）

URI是统一资源描述符，用于WWW的客户端。URI标志的是有关会话信息的外部的一个指针。这个字段是可选的。但是如果一旦出现，就必须先于第一个媒体轨出现，并且不能有多于一个。

5.6 电子邮件和电话号码（“e=”和“p=”）

e=<电子邮件地址>

p=<电话号码>

这两个字段描述了会议拥有者的联系信息。这个并不一定表示的是发起会话的那个人。这两个字段都是可选的。需要注意的是在上一版本的规范中，这两个字段必须出现一个，但是在这个版本中，这个限制去掉了。如果出现这两个字段，则必须先于第一个媒体轨出现。这两个字段都可以出现多次。

电话号码的格式应该以国际化的格式出现。如：

p=+1 617 555-6011

这两个字段都可以包含一个文本信息，通常表示人的名字。例如：

e=j.doe@example.com(Jane Doe)

也可以以以下形式出现：

e=Jane Doe<j.doe@example.com>

这两个字段的字体必须是ISO 10646或者是ISO-8859-1或者是字段“a=charset”表示的。

5.7 连接数据（“c=”）

c=<网络类型> <地址类型> <链接地址>

这个字段包含连接数据。

一个会话描述必须在每个媒体层都包含“c=”字段或者在会话层包含一个“c=”字段。如果这两个层都出现的话，则媒体层出现的“c=”会覆盖会话层出现的“c=”字段的值。

<网络类型>是一个文本字符串，最初只定义了“IN”，表示“Internet”，未来会定义其他值。

<地址类型>最初只定义了IP4和IP6。

<链接地址>标志连接的地址。后续可能会包含额外的字段，取决于<地址类型>字段。

当<地址类型>为IP4和IP6时，连接地址定义如下：

- 当会话是多播时，地址为多播组地址；当会话是单播时，地址为单播地址，并且为媒体数据的源地址。
- 如果地址类型是IPv4的，则还需要给出TTL的值，TTL表示包的生存时间，范围是

0~255。IPv6没有TTL

分层编码机制通常把媒体流分为多个层的媒体数据。接收方可以根据自己的能力或需要只接收其中的某些层。这种编码通常以多个多播组地址的方式发送。对于多个多播组的情况下，我们设计链接地址如下：

<组播地址基地址>[/<ttl>]/<地址数量>

如果地址数量没有给出来，则默认为1。例子如下：

c=IN IP4 224.2.1.1/127/3

表示224.2.1.1, 224.2.1.2, 和 224.2.1.3, TTL都是127。

这个在语法上等价于：

c=IN IP4 224.2.1.1/127

c=IN IP4 224.2.1.2/127

c=IN IP4 224.2.1.3/127

类似的，对于IPv6来说，如下：

c=IN IP6 FF15::101/3

等价于：

c=IN IP6 FF15::101

c=IN IP6 FF15::102

c=IN IP6 FF15::103

这种多个地址的表示法只能用于媒体层，而不能在会话层使用，而且也不能用于单播地址。

5.8 带宽（“b=”）

b=<带宽类型>:<带宽值>

这个字段的意思是本会话或者媒体需要占用的带宽。本规范制定了两种带宽类型，分别如下：

CT：这个带宽类型表示会话所占的所有的带宽的大小。当用于RTP会话时，表示所有的RTP会话所占用的带宽。

AS：这个带宽类型是针对特定应用的。通常，这表示某应用所占用的最大带宽。当用于RTP会话时，表示单一RTP会话所占用的带宽。

注意：**CT**表示的是所有媒体轨所占用的带宽，**AS**表示单一媒体轨所占用的带宽。

如果在<带宽类型>前加上“X-”前缀，则表示这些带宽类型只是在实验中使用，例如：

b=X-YZ:128

并不推荐“X-”前缀的用法，应该采用在IANA注册的形式来增加带宽类型。**SDP**解析器应该忽略不认识的带宽类型。带宽类型是字母+数字的形式，没有长度限制，但是通常很短。

<带宽值>的单位缺省是kilobits，在定义带宽类型的时候也会定义这个单位。

5.9 时间信息（“t=”）

t=<起始时间> <结束时间>

“t=”字段标志了会话的起始时间和结束时间。如果会话在不规则的多个时间段内有效，则可能会出现多个“t=”。如果时间段是规则的，则应该使用“r=”字段，后面会介绍。

第一个和第二个子字段分别表示起始时间和结束时间，时间格式采用NTP格式，为了转

换为UNIX时间，需要减去2208988800。NTP时戳为64位，时间可以表示到2036年。

如果<结束时间>是0，则表示会话时长是不受限的，不过需要在<起始时间>才有效。
如果<起始时间>是0，则表示会话是永恒的。

用户应该避免使用不受限或者永恒的会话时间，因为这些时间信息没意义，难以安排时序。

5.10 重复时间（“r=”）

r=<重复的间隔> <持续时长> <相对起始时间的偏移链表>

字段“r=”标志了会话的重复时间。例如：如果一个会话有效时间为每个星期一早上10点和星期二早上11点，每次持续1个小时，总共持续3个月。则字段“t=”的<起始时间>为第一个星期一的早上10点，<重复间隔>为1个星期，<有效时长>为1个小时，<相对起始时间的偏移链表>为0和25小时。字段“t=”的<结束时间>为三个月后。缺省情况下，单位都为s，因此“r=”和“t=”字段如下：

t=3034423619 3042462419

r=604800 3600 0 90000

通常，为了更紧凑的表示，也可以使用天，小时，分钟为单位。例如：r=7d 1h 0 25h。

5.11 时域（“z=”）

为了在夏令时和标准时间之间转换，有必要需要知道对基准时间的偏移量。

为了表示一个会话在夏天和冬天一个时间，有必要指出是哪个时域的。“z=”字段允许列举一系列基于基准时间的调准时刻及其偏移量。如：

z=2882844526 -1h 2898848070 0

这表示在时刻2882844526，基准时刻需要往前调1个小时，在时刻2898848070，则不需要调准。

5.12 加密密钥（“k=”）

k=<方法>

k=<方法>:<密钥>

如果在一个安全信道上进行传输，则可以通过SDP来传达密钥。最简单的方式是采用字段“k=”来协商。注意，现在这种方式已经是不推荐的了。

字段“k=”可以在所有媒体轨之前，这表示应用于所有的媒体轨，或者单独用于某个媒体轨。本规范定义了如下几种方法：

- k=clear:<密钥>

在这种方法中，密钥是不经过任何转换的。除非保证传输通道绝对安全，否则不允许使用这种方法。

- k=base64:<编码后的密钥>

在这种方法中，密钥经过base64的编码。除非保证传输通道绝对安全，否则不允许使用这种方法。

- k=uri:<URI，通过此来获得密钥>

在这种方法中，给出一个URI。通过这个URI，可以获得密钥，这中间可能需要经过某些认证。

- **k=prompt**

在这种方法中，没有给出密钥。但是使用这种方法的会话是经过加密的。用户加入会话时会提示输入密钥。不过这种使用用户特定密钥的方式也是不推荐的。

再次强调：除非保证SDP在一个安全的通道上传输，否则是不会使用“k=”字段的。

5.13 属性（“a=”）

a=<属性>

a=<属性>:<值>

属性是扩展SDP的主要方式，有会话层属性和媒体层属性。会话层的属性应用于所有的媒体轨，媒体层的属性只应用于当前的媒体轨。

属性有两种方式：

- 特性属性，以a=<特性>表示，例如：“a=recvonly”
- 值属性，以a=<属性>:<值>表示，例如“a=orient:landscape”

属性的名字必须使用US-ASCII子集。属性的值是字符串，除了0x00,0x0a,0x0d这三个字符除外。缺省情况下，会使用UTF-8的ISO-10646字体集，这个字段是唯一不受“a=charset”影响的。当然，不排除会特别设计采用“a=charset”的字体集的扩展字段。

属性必须在IANA注册，如果解析器碰到不认识的属性字段，则应该忽略它。

5.14 媒体描述（“m=”）

m=<媒体类型> <端口> <协议> <格式类型>

一个会话描述可能包含多个媒体描述。每个媒体描述都是以“m=”字段开始的，结束于下一个“m=”或者到整个会话描述结束。

<媒体类型>，本规范只定义了视频、音频、文本、应用、消息这几种类型。

<端口>，被发送媒体流的端口。这个字段的意义依赖于“c=”字段和<协议>字段。如果不使用连续的端口或者没按照偶数RTP端口、奇数RTCP端口，则必须使用“a=rtcp:”来标识RTCP的端口。

对于分层编码的码流，如果发送采用单播地址，则必须使用端口来区分这些码流。如下：

m=<媒体类型> <端口>/<端口数量> <协议> <格式类型>

在这种情况下，端口的意义依赖传输协议。对于RTP协议来说，缺省情况下只采用偶数端口来发送RTP数据，对应的端口加1来发送RTCP数据。

如果在“c=”字段中采用多地址，在“m=”字段中采用多端口，则，这些地址和端口会默认为一一对应。

对于不同的媒体轨使用相同的端口的情况，本规范没有进行定义。需要在某些具体应用中去自己定义。

<协议>指传输协议。传输协议的意义以来与“c=”字段。例如“c=”字段中的IP4表示在IP4上的协议。本规范定义如下协议：

- udp: udp上的一个未指定协议
- RTP/AVP: 指RTP协议
- RTP/SAVP: 指SRTP协议

<格式类型>，媒体格式的描述。这个参数可能是一个链表，表示多个媒体格式的类型。这个字段的意义依赖于<协议>字段。

如果<协议>字段是"RTP/AVP" 或者"RTP/SAVP"，则媒体格式表示RTP负载格式的编号。当出现的是一个链表的时候，表示链表中的媒体格式都可以用于当前的媒体轨，不过第一个媒体格式为缺省的格式。"a=rtpmap:"属性是用来动态匹配媒体格式编号和媒体格式的。"a=fmtp:"属性可能用来描述媒体格式具体的参数。

如果<协议>字段是udp，媒体格式指定了媒体类型为音频、视频、文本、应用或者消息。这些媒体类型定义了对应的UDP传输的包格式。对于使用其他传输协议的，<媒体格式>这个字段是在相应的传输协议中定义。

6 SDP 属性

下面是本规范定义的属性。这些属性并不是详尽的，因为有可能应用会加入他们自己需要的另外的属性。在8.2.4部分描述了新属性注册的流程。

- **a=cat:<category>** 分类属性

这个属性给出了用圆点分隔的回话的分层分类。分类是没有预先注册好的。这是一个会话层的属性，并且它不依赖于字符集。

- **a=keywds:<keywords>** 关键字属性

类似分类属性，这也是辅助接受者识别期望的属性。这个属性允许接受者可以用来选择与关键字中描述的信息相符合的会话。这些关键字是没有预先注册好的。这是一个会话层的属性。它依赖于字符集。也就是说如果有定义字符集，则它的语义符合字符集定义的语义；如果没有定义字符集，则默认为ISO 10646/UTF-8。

- **a=tool:<name and version of tool>** 工具集属性

这个属性给出了创建会话描述的工具的名字和版本号。这是一个会话层的属性，不依赖于字符集。

- **a=ptime:<packet time>** 包的时长

这个属性给出了每个媒体包的时长，以毫秒为单位。这通常只用于音频数据，当然，如果在其他情况下也有意义，也可以使用。对于解码来说，这不是一个必要的属性，它通常作为编码/打包的时长的建议。这是一个媒体层的属性，不依赖于字符集。

- **a=maxptime:<maximum packet time>** 最大的包时长

这给出了每个媒体包的最大的时长，以毫秒为单位。如果是基于帧的编码，则这个通常是帧长的整数倍。这通常只用于音频数据，当然，如果在其他情况下也有意义，也可以使用。这是一个媒体层的属性，不依赖于字符集。注意，这个属性是在RFC2327后才提出来的，如果实现没有更新，则会忽略该属性。

- **a=rtpmap:<payload type> <encoding name>/<clock rate> [/<encoding parameters>]** 媒体负载匹配属性

这个属性是用来匹配RTP负载编号和编码负载格式。它也提供时钟频率和编码参数。这是一个媒体层的属性，不依赖于字符集。

尽管某些规范定义了负载编号和负载编码的静态匹配，但是更普遍的是使用"a=rtpmap:"来动态匹配。静态匹配的例子是8KHz采样率的单通道的u-law的PCM编码。这在RFC3551中定义，负载编号为0。这样的媒体流的描述如下表示：

m=audio 49232 RTP/AVP 0

动态匹配的例子是16KHz采样率的立体音的16bit线性编码。如果我们想用编号为98表

示这个媒体流，则描述如下：

m=audio 49232 RTP/AVP 98

a=rtpmap:98 L16/16000/2

每个rtpmap属性可以表示一个媒体流，例如：

m=audio 49230 RTP/AVP 96 97 98

a=rtpmap:96 L8/8000

a=rtpmap:97 L16/8000

a=rtpmap:98 L16/11025/2

如果RTP的规范定义动态匹配的类型，它必须定义有效的编码名字集合，如果使用SDP来描述，还要定义注册的方式。协议“RTP/AVP”和“RTP/SAVP”采用媒体子类型来定义编码名字，例如使用“audio/l8”或者“audio/l16”。

在音频流中，<编码参数>表示音频通道的数量。这个参数是可选的，如果通道数为1，也可以忽略。

在视频流中，没有定义编码参数。

将来可能会定义新的编码参数，但是不应该定义与编解码相关的参数。在rtpmap中定义的参数是那种为了方便选择会话的参数，如果是与编解码相关的参数，应该定义在其他的属性如“a=fmtp:”中。

注意：RTP音频格式通常没有每个包中包含的采样的数量。如果没有采用缺省的打包格式，则使用“ptime”属性表示上述意思。

- **a=recvonly**只接受模式

这个属性标志工具应使用只接受模式。这个属性可能属于会话层，也可能属于媒体层，并且不依赖于字符集。注意，recvonly只用于媒体，不用于控制协议。

- **a=sendrecv**发送接受模式

这个属性标志工具应使用发送接受模式。交互式会议中使用缺省模式为只接受模式的工具时，这个属性是必须的。这个属性可能属于会话层，也可能属于媒体层，并且不依赖于字符集。

- **a=sendonly**只发送模式

这个属性标志工具应使用只发送模式。在某些情况下，双方通信可能一方只发送，一方只接受，这个时候就使用这个参数。这个属性可能属于会话层，也可能属于媒体层，并且不依赖于字符集。注意，recvonly只用于媒体，不用于控制协议。

- **a=inactive**非激活模式

这个标志工具应使用非激活模式。在交互式会议中，当一个用户可以让另一个用户等候的时候，必须使用这个属性。在一个非激活的媒体流中，没有媒体数据被发送。注意，RTCP包仍然会发送。这个属性可能属于会话层，也可能属于媒体层，并且不依赖于字符集。

- **a=orient:<orientation>**方位属性

这个属性只用于白板工具，标志了工作区在屏幕的位置。这是一个媒体层的属性。允许的值是：“portrait”，“landscape”，和“seascape”。不依赖于字符集。

- **a=type:<conference type>**会议类型

这个属性标志了会议的类型。建议的值是：“broadcast”，“meeting”，“moderated”，“test”，和“H332”。对于“type:broadcast”类型，“recvonly”是缺省的模式，“type:meeting”暗示使用“sendrecv”等等。这是一个会话层的属性，不依赖于字符集。

- **a=charset:<character set>**字符集属性

这个属性标志了会话名字和会话信息使用的字符集。缺省是ISO-10646/UTF-8字符集。其他字符集如下表示，如使用ISO 8859-1字符集：a=charset:ISO-8859-1。

这个属性是一个会话层的属性，不依赖于字符集。字符集必须是已经在IANA注册的字符集。这个属性的值采用US-ASCII字符串，大小写敏感的。如果实现不认识这个字符集，则会忽略这个值，并把它影响的地方都当成是字符串处理。

注意，一个字符集禁止使用0x00，0x0a，0x0d这三个字符。如果使用，则会采用其他形式（如转义字符）表示，避免与每行的结束符混淆。

- **a=sdplang:<language tag>SDP语言**

这个属性可能是会话层或者媒体层的。作为一个会话层属性，表示所有的媒体层的SDP信息都是被这个属性的值所影响。作为一个媒体层属性，只会影响那个媒体层的信息。这个属性肯能出现多个，最重要的最先出现，最不重要的最后出现。

通常情况下，不推荐包含多个语言属性。取而代之的是发送多个SDP，每个SDP采用一种语言。但是，这有时候不可能实现，因此才会允许之前那种发送多个语言属性的情况。

“sdplang”属性的值必须是RFC3066中的某个语言标志，以US-ASCII形式出现。

- **a=lang:<language tag>语言标志**

这个属性可能是会话层或者媒体层的。作为一个会话层属性，表示所有的媒体层的SDP信息都是被这个属性的值所影响。作为一个媒体层属性，只会影响那个媒体层的信息。这个属性肯能出现多个，最重要的最先出现，最不重要的最后出现。

“lang”属性的值必须是RFC3066中的某个语言标志，以US-ASCII形式出现。

- **a=framerate:<frame rate>帧率**

这个属性给出了视频帧的最大帧率。这通常作为编码的推荐参数。如果是小数，则以下面的形式表示：<integer>.<fraction>。这是一个媒体层属性，只用于视频媒体流，不依赖于字符集。

- **a=quality:<quality>质量**

这个属性给出了编码的质量要求，以整数形式出现。这个参数给出了帧率和图像质量的权衡。对于视频而言，值的范围在0~10之间，推荐值如下：

- 10：最好质量。
- 5：缺省。
- 0：质量最差。

这是一个媒体层属性，不依赖于字符集。

- **a=fmtp:<format> <format specific parameters>编解码参数**

这个属性标志了具体格式的编解码参数，SDP本身并不需要明白这些参数。媒体格式必须是“m=”中标识的某个媒体类型。编解码参数可能是采用SDP传输的任意参数，原封不动透传给解码器。每个媒体类型最多只能有一个这种属性。

这是一个媒体层属性，不依赖于字符集。

7 安全考虑

SDP经常用于单播会话的SIP中，使用offer/answer模式协商参数。在这种方式下，这些协议的安全需要考虑。

SDP是描述多媒体会话的一个会话描述格式。不过接收端必须有下面这个意识：SDP并不是可信的，除非SDP通过某个已被认证的通道进行传输。SDP可以通过多种传输协议进行分发，而各种传输协议的认证的特点又是各种各样的。对于某些传输协议，根本就没有安全的保护。因此，如果SDP不是通过可信的通道传输的，终端需要小心的执行。因为在某些攻击下，SDP已经不是期望的那个，因此，目的地址也已经不是期望的那个，甚至媒体的安全

可能受到侵害。这时候，做的任何操作都取决于终端。

分发SDP的其中一个协议是SAP。SAP本身提供了加密和认证的机制，但是由于SAP的特性，某些情况下，认证机制使用不了。例如发起者事先就不被接受者知道或者公共密钥无法提供。

当通过某个没被认证或者不可信的源接收到SDP时，终端需要采取某些预防措施。SDP可能包含一些信息，而终端需要根据这些信息启动一些软件。终端只能启动那些被配置好的，不可以启动其他的软件。通常终端不应该解析SDP后直接启动软件，而应该通知用户，然后得到用户的准许等候才能启动软件。因此，除非得到用户授权，终端不能直接加入某个交互的多媒体会话。由于不是很容易知道会话是否是交互的，因此终端可以默认会话就是交互的。

在本规范中没有任何属性允许SDP的接受者以默认的传送方式启动多媒体工具。在某些情况下，可以定义这种属性。如果定义了这种属性，终端必须忽略它或者通知用户加入会话会导致媒体数据自动传输。对于不认识的苏醒，默认的行为是忽略它。

在特定的环境下，通常会由中间层系统来拦截并解析信令协议中的SDP。这通常包含但不仅限于以下目的：让媒体数据得以穿透防火墙，或者标志并区分优先顺序，并选择性的阻挡媒体数据。某些情况下，这些中间件会修改SDP。这通常是不推荐的，除非SDP以下面这种方式传输：允许中间件确认SDP的真实性并且SDP的源可以建立适合当前情况下的SDP。SDP本身并不包含任何信息允许这种检查，这种检查以来封装协议（如SIP或者RTSP）。

采用“k=”字段会造成严重的安全隐患，因为它以明文的方式传输密钥。除非在私有或者可信的通道上传输SDP，否则不允许使用“k=”字段传输密钥。而且，“k=”字段也无法标识或者协商加密算法。因为它是一种对称的密钥，而不是针对私密性和完整性提供独立的密钥，因此，它的用途非常有限。总之，不推荐使用“k=”字段。

8 IANA 考虑

8.1 “application/sdp”媒体类型

这是一个在RFC2327中定义的媒体类型，在本规范中会有相应的更新，如下所示：

收件人：ietf-types@iana.org

主题：媒体类型“application/sdp”的注册

类型名字：application

子类型名字：sdp

必选参数：无

可选参数：无

编码考虑：SDP文件主要是UTF-8文本格式。可以使用“a=charset”标识其他的字符集。在SDP中不能直接描述二进制内容。

安全考虑：参考RFC4566的第七部分

互用性考虑：参考RFC4566

发布规范：参考RFC4566

使用这种媒体类型的应用场景：IP通话，视频会议，流媒体，即时通讯等。

8.2 参数的注册

本规范描述了7个可能在IANA注册的字段。在SDP规范中使用BNF，他们是："media", "proto", "fmt", "att-field", "bwtype", "nettype"和"addrtype"。

8.2.1 媒体类型

媒体类型的集合很小，而且只在很少的情况下可以扩展。这种规则同样应用于顶层的媒体内容类型。当需要定义新的类型时，需要提供一个标准的RFC跟踪，表述为什么当前的类型都不满足。

本规范组侧了下面的媒体类型："audio", "video", "test", "application", 和 "message"。

注意：媒体类型"control"和"data"在之前的规范中定义，然而他们的语义没有被明确说明而且很少使用，因此他们在本规范中被去掉了。如果这些类型在将来被使用，则必须提供标准的RFC跟踪。

8.2.2 传输协议

"proto"字段描述使用的传输协议。这个应该参考某个标准的RFC跟踪协议。本规范注册三个值："RTP/AVP", "RTP/SAVP", "udp"。

新的传输协议应该在IANA注册。注册必须提供某个RFC描述此协议，尽管这个协议可能并没成为正式的协议。而且，协议的注册同样需要描述"fmt"的命名空间。

8.2.3 媒体格式

每个"proto"字段定义的传输协议都有一个与之关联的"fmt"命名空间，用来描述在此协议上传输的媒体类型。这些媒体格式包括了所有可能的编码格式。

"RTP/AVP"和"RTP/SAVP"定义的RTP负载格式必须使用负载类型编号作为他们的"fmt"值。如果负载编号是动态匹配的，则还需要加上"rtptime"指定具体的格式名字及参数。建议其他的RTP传输协议指定他们的"fmt"命名空间。

对于"udp"协议，应该注册新的格式。建议使用已有的媒体子类型。如果没有，则建议按照IETF的规程注册一个。

对于其他协议，应该根据定义"proto"的规范中的规则进行注册。注册一个新的媒体格式必须指定属于哪个传输协议。

8.2.4 属性字段名字

属性名字必须在IANA注册，因为有可能和已有的名字冲突。不认识的属性可以忽略，但是冲突的属性会引起严重的问题。

新的属性注册会根据RFC2334进行，需要提供下面信息：

- 联系名字，邮件地址，电话号码
- 属性名字
- 英语中的完整形式属性名字

- 属性的类型，会话层的，还是媒体层的，还是两层都有
- 是否受字符集影响
- 属性的用途
- 一个描述这个属性的值的规范

上面的信息只是一个最小集。如果想被广泛使用并且具有互用性，则应该提供一个标准的RFC跟踪描述此属性。

注册的提供者必须保证此属性符合SDP属性的本质，最显著的是平台无关性。

现有的已注册的属性如下所示：

Name	Session or Media level?	Dependent on charset?
cat	Session	No
keywds	Session	Yes
tool	Session	No
ptime	Media	No
maxptime	Media	No
rtpmap	Media	No
recvonly	Either	No
sendrecv	Either	No
sendonly	Either	No
inactive	Either	No
orient	Media	No
type	Session	No
charset	Session	No
sdplang	Either	No
lang	Either	No
framerate	Media	No
quality	Media	No
fntp	Media	No

8.2.5 带宽说明

本规范强烈不推荐进行带宽类型的扩展。新的带宽类型必须在IANA注册。必须提供标准的RFC跟踪，并解析为何当前的带宽类型无法满足。

本规范定义的带宽类型是 "CT" 和 "AS" 。

8.2.6 网络类型

如果SDP需要在非互联网的环境下使用，则可能需要注册新的网络类型。网络类型的集合应该很小，并且很少会扩展。一个新的网络类型的注册，必须先注册至少一个地址类型。一个新的网络类型的注册必须参考某个RFC，在这个RFC中描述了网络类型和地址类型的细节，并描述他们是怎么使用的。

在本规范中，IANA已经注册了网络类型"IN"表示互联网。

8.2.7 地址类型

新的地址类型可能会在IANA上注册。一个地址类型只在某个特定的网络类型的背景下有意义。一个地址类型的注册必须指定属于哪个网络类型。一个新的地址类型的注册必须参

考某个RFC，在RFC中描述了地址类型的语法。通常，并不频繁的注册地址类型。
在本规范中，IANA注册了地址类型"IP4"和"IP6"。

8.2.8 注册流程

为了注册某个新的字段，作者必须提供下面的信息：

- 联系名字，邮件地址，电话号码
- 注册的名字
- 英文的完整形式
- 名字类型（"media", "proto", "fmt", "bwtype", "nettype", 或者 "addrtype"）。
- 用途的描述
- 参考的规范

IANA可能把注册提交IESG审核，还可能会进行修订。

8.3 加密密钥和方法的匹配

IANA之前还提供了一个表格匹配SDP加密密钥和方法之间的匹配。但是这个表格被删除了，原因是" k="是不可扩展的。同时也不支持任何注册。

9 与 RFC2327 的不同的总结

本规范的结构已经发生变化，根据使用包含了大量的规范的说明。本规范在大体上与RFC2327兼容。除了以下几点：

- 第九部分的ABNF语法已经彻底的修订和更新，改正了一些错误，并且把一只的一些不一致修改掉了。
- 本规范还包含如何在IANA上进行属性或者其他参数的注册流程。注意，"text"和"message"是有效的媒体类型，而"control"和"data"已经是无效的媒体类型。
- 本规范还使用了RFC2119中定义的要求级别。某些要求级别比RFC2327中要严格。
- 本规范还注册了"RTP/SAVP"协议的RTP规范及其"fmt"命名空间。
- 加入了属性 "a=inactive" 和 "a=maxptime"。
- RFC2327中规定"e="或者"p="必须包含一个，本规范中两者都是可选的。
- 还有一个重要的是本规范中描述了" k="的限制，并强烈不推荐使用。
- 不推荐使用"x-"前缀的参数。