



# 中华人民共和国国家标准

GB/T XXXXX.9.1—YYYY

---

信息技术 先进音视频编码  
第 8 部分：IP 网络传输  
第 1 子部分：移动视频 RTP 净载格式  
Information technology - Advanced coding  
of audio and video  
Part 8: AVS over IP  
Sub-part 1: AVS-M RTP Payload Format  
(送审稿)

2005 年 11 月

××××-××-××发布

××××-××-××实施

---

国家质量监督检验检疫总局 发布

目 次

前 言..... III

引 言..... IV

1 范围..... 5

2 规范性引用文件..... 5

3 术语和定义..... 5

3.1 定义..... 5

3.2 缩写..... 5

4 RTP 净载格式..... 6

4.1 RTP 头的使用..... 6

4.2 RTP 净载格式的公共结构..... 7

4.3 NALU 头的使用..... 7

4.4 打包模式..... 8

4.5 解码顺序号（DON）..... 8

4.6 单 NALU 包..... 10

4.7 复合包..... 10

4.7.1 单时间复合包..... 11

4.7.2 多时间复合包..... 14

4.8 分割单元..... 16

5 打包规则..... 18

5.1 公共打包规则..... 18

5.2 单 NALU 模式..... 19

5.3 非交织模式..... 19

5.4 交织模式..... 19

6 净载格式参数..... 19

6.1 MIME 注册..... 19

6.1.1 可选 MIME 参数的语义..... 19

6.2 SDP 参数..... 23

6.2.1 MIME 参数与 SDP 域之间的匹配..... 23

6.2.2 在 SDP 提供/回答模型中的用法..... 24

6.2.3 在申明性的会话描述中的用法..... 26

6.3 参数集考虑..... 26

7 安全考虑..... 28

8 拥塞控制..... 28

9 IANA 注册..... 28

10 解包过程（资料性）..... 28

10.1 单 NALU 和非交织模式..... 28

10.2 交织模式..... 29

10.2.1 解交织缓冲区尺寸..... 29

10.2.2 解交织过程..... 29

10.3 额外的解包指导原则..... 30

## 前 言

GB/T XXXXX在《信息技术 先进音视频编码》的总标题下，包括以下九个部分：

- 第1部分：系统；
- 第2部分：视频；
- 第3部分：音频；
- 第4部分：一致性测试；
- 第5部分：参考软件；
- 第6部分：数字媒体版权管理；
- 第7部分：移动视频；
- 第8部分：IP网络传输；
- 第9部分：文件格式。

本部分为GB/T XXXXX的第8部分的第1子部分：移动视频RTP净载格式。

本部分由中华人民共和国信息产业部提出。

本部分由全国信息技术标准化技术委员会归口。

本部分由 XX 起草。

本部分主要起草人：XX。

## 引 言

本部分定义 AVS-M 视频的 RTP 净载格式。该 RTP 净载格式允许在每个 RTP 净载中包含一个或多个 AVS-M 视频 NAL 单元（简称 NALU）。该 RTP 净载格式应用范围宽广，支持简单的低比特率对话、采用交织传输的网络流媒体以及高比特率的视频点播。

信息技术 先进音视频编码 第 8 部分:IP 网络传输  
第 1 子部分: 移动视频 RTP 净载格式

1 范围

本部分规定的AVS-M视频RTP净载格式只适用于AVS-M视频NAL单元流，而不适用于AVS-M视频标准附录D所定义的字节流格式。因此，本部分的应用范围包多媒体对话、视频电话或视频会议、流媒体和IP电视等。

本部分定义了一些新的NAL单元类型，这些NAL单元类型在AVS-M视频标准中的语义为“未规定”。此外，本部分还扩展了NAL单元中forbidden\_zero\_bit和nal\_ref\_idc的语义。

2 规范性引用文件

下列文件中的条款通过本部分的引用而成为本部分的条款。凡是注日期的引用文件，其随后所有的修改单（不包括勘误的内容）或修订版均不适用于本部分，然而，鼓励根据本部分达成协议的各方研究是否可使用这些文件的最新版本。凡是不注日期的引用文件，其最新版本适用于本部分。

- GB/T XXXXX.7—YYYY 信息技术 先进音视频编码 第7部分:移动视频（AVS-M）
- IETF RFC 3550, RTP: a transport protocol for real-time applications
- IETF RFC 2327, SDP: session description protocol
- IETF RFC 3264, An offer/answer model with session description protocol (SDP)
- IETF RFC 3584, The base16, base32, and base64 data encodings

3 术语和定义

AVS-M移动视频标准中的定义和缩写适用于本部分。另外，下面的定义和缩写也适用于本部分。

3.1 定义

**解码顺序号（DON）**

净载结构中的一个域或一个导出变量，指定NALU**解码顺序**。取值范围为0..65535。当达到最大值后，下一个DON值为0。

**NALU解码顺序**

符合AVS-M视频标准第7.2.1.1小节规定的NAL单元的顺序。

**传输顺序**

按RTP序列号（RTP sequence nubmer）上升顺序（在模算术的情况下）的包顺序。

**具有媒体知识的网络单元（media aware network element, MANE）**

能够解析RTP净载头或净载的某些部分并针对内容作出相应处理的网络单元，如应用层网关。

注：与普通路由器或网关相比，MANE必须知道信号的内容（如媒体流净载类型的匹配），另外当采用SRTP时MANE必须是可信的。MANE的优点是可以根据媒体编解码的需要进行丢包。例如，如果由于某一链接的拥塞MANE必须丢包，它可以识别并丢弃那些对用户体验影响最小的包。

3.2 缩写

- DON: 解码顺序号 (Decoding Order Number)
- DONB: 解码顺序号基数 (Decoding Order Number Base)
- DOND: 解码顺序号偏移 (Decoding Order Number Difference)
- FEC: 前向纠错 (Forward Error Correction)

- FU: 分割单元 (Fragmentation Unit)
- GDR: 逐渐解码刷新 (Gradual Decoding Refresh)
- HRD: 虚拟参考解码器 (Hypothetical Reference Decoder)
- IDR: 即时解码刷新 (Instantaneous Decoding Refresh)
- IEC: 国际电子技术委员会 (International Electrotechnical Commission)
- ISO: 国际标准化组织 (International Organization for Standardization)
- ITU-T: 国际电信联盟电信标准化部门 (International Telecommunication Union, Telecommunication Standardization Sector)
- MANE: 具有媒体知识的网络单元 (Media Aware Network Element)
- MIME: 多用途因特网邮件扩展 (Multipurpose Internet Mail Extensions)
- MTAP: 多时间复合包 (Multi-Time Aggregation Packet)
- MTAP16: 16位时间戳偏移的多时间复合包 (MTAP with 16-bit timestamp offset)
- MTAP24: 24位时间戳偏移的多时间复合包 (MTAP with 24-bit timestamp offset)
- MTU: 最大传输单元 (Maximum Transfer Unit)
- NAL: 网络抽象层 (Network Abstraction Layer)
- NALU: NAL单元 (NAL Unit)
- PPS: 图像参数集 (Picture Parameter Set)
- RTP: 实时传输协议 (Real-time Transport Protocol)
- SEI: 辅助增强信息 (Supplemental Enhancement Information)
- SPS: 序列参数集 (Sequence Parameter Set)
- STAP: 单时间复合包 (Single-Time Aggregation Packet)
- STAP-A: A类单时间复合包 (STAP type A)
- STAP-B: B类单时间复合包 (STAP type B)
- TS: 时间戳 (Timestamp)

4 RTP 净载格式

4.1 RTP 头的使用

IETF RFC 3550中定义了RTP头的格式，如图1所示。本净载格式对其中数据域的使用方法符合IETF RFC 3550的规定。

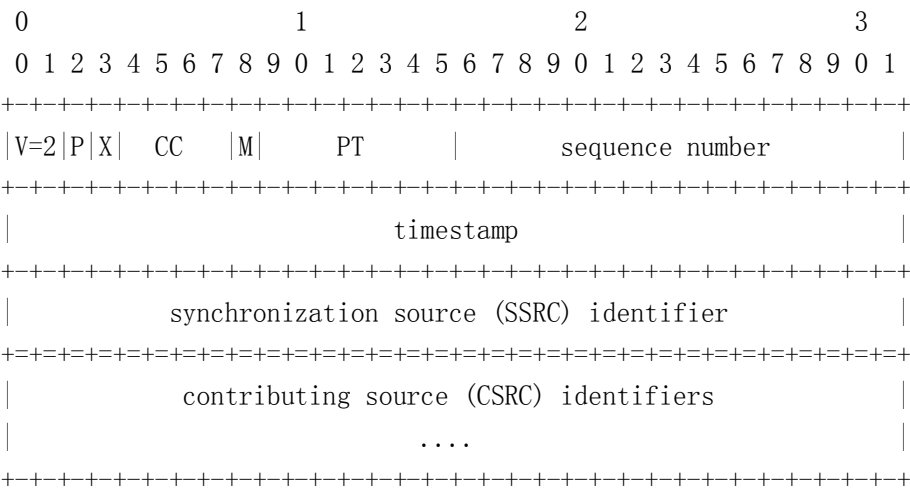


图1. RTP头

如果每个RTP包中封装一个NALU，推荐使用第4.6小节规定的RTP净载格式。复合包和分割单元的RTP净载格式以及RTP头使用方法分别在4.7和4.8小节给出。

本RTP净载格式对RTP头信息的设定如下（未提到的数据域的设定由IETF RFC 3550规定）：

**标记位（M）：**1位

一个访问单元（由RTP时间戳指定）的最后一个包的标记位设为1，这个访问单元其它包的标记位均设为0。对于复合包（STAP或MTAP），标记位的值应等于当该包最后的NALU单独打包时的标记位值。RTP包接收器可以通过这个比特得知是否已经收到访问单元的最后一个包，但是不能完全依靠这个信息。

注：包含多个NALU的复合包只有一个标记位。如果网关将一个复合包重新打包成多个包，这些包的标记位不能可靠设定。

**时间戳（Timestamp）：**32位

RTP时间戳设为内容的采样时间戳，所用的时钟必须为90000赫兹，即时间单位为1/90000秒。如果NALU本身不具有时间属性，如参数集和SEI NALU，RTP时间戳设为包含该NALU的访问单元中的编码图像的RTP时间戳。MTAP的RTP时间戳的设定在4.7.2小节规定。

4.2 RTP 净载格式的公共结构

本净载格式定义三个不同的净载结构。接收器可以通过RTP净载的第一个字节识别净载结构。RTP净载的第一个字节同时也作为RTP净载头的一部分。这个字节总是NALU的头，其中的NALU类型指定净载结构。可能的结构如下：

单NALU包：净载中仅包含一个NALU。NALU头中的NALU类型与比特流中的原始值相同，取值范围为0..23。见4.6节。

复合包：净载中包含多于一个NALU。复合包分为四类，即A类单时间复合包（STAP-A）、B类单时间复合包（STAP-B）、16位时间戳偏移的多时间复合包（MTAP with 16-bit timestamp offset）和24位时间戳偏移的多时间复合包（MTAP with 24-bit timestamp offset），对应的NALU类型值分别为24、25、26和27。见4.7节。

分割单元：用于将一个NALU分割到多个RTP包中。分割单元分为两类，A类分割单元（FU-A）和B类分割单元（FU-B），对应的NALU类型值分别为28和29。见4.8节。

表1给出了不同NALU类型值对应的净载结构。

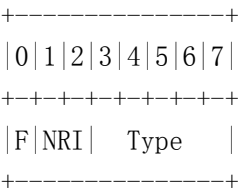
表1. NALU类型值对应的净载结构

类型	净载结构	净载结构全名	章节
0	未规定		—
1-23	单NALU	单NAL单元包	4.6
24	STAP-A	A类单时间复合包	4.7.1
25	STAP-B	B类单时间复合包	4.7.1
26	MTAP16	16位时间戳偏移的多时间复合包	4.7.2
27	MTAP24	24位时间戳偏移的多时间复合包	4.7.2
28	FU-A	A类分割单元	4.8
29	FU-B	B类分割单元	4.8
30-31	未规定		—

4.3 NALU 头的使用

NALU头的结构为：





其中F、NRI和Type分别对应AVS-M视频标准中的语法元素forbidden\_zero\_bit、nal\_ref\_idc和nal\_unit\_type。本小节给出本部分中F和NRI的语义，Type的语义见表1。

F：1位

值0表示NALU中不宜包含比特错误或其它语法错误。值1表示NALU头和净载中可能包含比特错误或其它语法错误。

如果检测到NALU中有比特错误，MANE宜将F设为1。AVS-M标准要求F为0。对于F为1的NALU，解码器最简单的操作是将其丢弃，然后进行错误掩蔽。

NRI：2位

除了必须符合AVS-M视频标准中的规定之外，NRI值表明相对传输优先级。MANE可以利用这个信息更好地保护更重要的NALU。优先级由高到低的NRI值依次为11b、10b、01b和00b。

序列参数集、图像参数集和IDR图像对应的NALU中的NRI值宜为11b。

对NALU类型为24到29的NRI值的定义见4.7和4.8节。

4.4 打包模式

本部分规定了三种打包模式：单NALU模式、非交织模式（non-interleaved mode）和交织模式（interleaved mode）。在单NALU模式或非交织模式下，NALU按照解码顺序传输；交织模式允许NALU不按照解码顺序传输。

采用的打包模式决定了RTP净载中允许包含哪些NALU类型，打包模式的传送可以通过可选的MIME参数packetization-mode或外部方式。表2总结了各打包模式下允许的NALU类型，其中标为“未规定”的NALU类型被保留以便将来的扩展。发送端不宜发送这些NALU类型，接收端应忽略这些NALU类型。打包模式在第5节详细介绍。

表2. 各打包模式下允许的NALU类型（yes = 允许，no = 不允许， ig = 忽略）

NALU 类型	包	单NALU 模式	非交织 模式	交织 模式
0	未规定	ig	ig	ig
1-23	NALU	yes	yes	no
24	STAP-A	no	yes	no
25	STAP-B	no	no	yes
26	MTAP16	no	no	yes
27	MTAP24	no	no	yes
28	FU-A	no	yes	yes
29	FU-B	no	no	yes
30-31	未规定	ig	ig	ig

4.5 解码顺序号（DON）

在交织打包模式下，NALU的传输顺序允许与解码顺序不同。解码顺序号（DON）是净载结构中的一个域或一个导出变量，指定NALU解码顺序。

通过可选的MIME参数sprop-interleaving-depth控制传输顺序和解码顺序之间的耦合。当sprop-interleaving-depth值为0时，或按非解码顺序传输NALU被外部方式禁止时，则NALU必须按照解码顺序传输。当sprop-interleaving-depth值大于0时，或当外部方法允许NALU按非解码顺序传输时，则：

- MTAP16和MTAP24中的NALU顺序不必为解码顺序；
- 通过分解连续两个包得到的STAP-B、MTAP、和FU中的NALU顺序不必为解码顺序。

单NALU包、STAP-A和FU-A的RTP净载结构中不包括DON；STAP-B和FU-B的RTP净载结构中包括DON；MATP的RTP净载结构允许导出DON，见4.7.2小节。

注：如果一个FU-A出现在交织模式下，则该FU-A总是在一个FU-B之后。

注：如果发送端希望将每个NALU打成一个包并且不按照解码顺序传输，可以使用STAP-B包。

在单NALU打包模式下，NALU的传输顺序（由RTP序列号决定）必须与解码顺序相同。在非交织打包模式下，单NALU包、STAP-A和FU-A中的NALU的传输顺序必须与解码顺序相同。包含在STAP中的NALU必须按解码顺序出现。这样，解码顺序的提供首先通过STAP中的隐含顺序，然后通过确定STAP、FU和单NALU包顺序的RTP序列号。

得到包含在STAP-B、MTAP和由一个FU-B开始的一系列FU中的NALU的DON值的方法分别在4.7.1、4.7.2和4.8小节介绍。第一个（传输顺序）NALU的DON值可以任意设定。DON的取值范围为0..65535。当达到最大值后，下一个DON值为0。

确定包含在STAP-B、MTAP或由一个FU-B开始的一系列FU中的两个NALU的解码顺序的方法如下。

令DON(i)为传输顺序索引为i的NALU的DON值，规定函数don\_diff(m, n)如下：

If DON(m) == DON(n), don\_diff(m, n) = 0

If (DON(m) < DON(n) and DON(n) - DON(m) < 32768),  
don\_diff(m, n) = DON(n) - DON(m)

If (DON(m) > DON(n) and DON(m) - DON(n) >= 32768),  
don\_diff(m, n) = 65536 - DON(m) + DON(n)

If (DON(m) < DON(n) and DON(n) - DON(m) >= 32768),  
don\_diff(m, n) = - (DON(m) + 65536 - DON(n))

If (DON(m) > DON(n) and DON(m) - DON(n) < 32768),  
don\_diff(m, n) = - (DON(m) - DON(n))

若don\_diff(m, n)等于0则两个NALU的解码顺序可以为任一顺序，值为正表明传输顺序索引为m的NALU的解码顺序在传输顺序索引为n的NALU之前，值为负则与值为正的情形相反。

与DON有关的域（DON、DONB和DOND，见4.7节）的值必须符合以下条件：由DON值按上述方法决定的解码顺序必须与NALU解码顺序相同。如果两个NALU的解码顺序互换后的新顺序与NALU解码顺序不一致，则它们必须具有不同的DON值。如果NAL单元流中两个连续的NALU的顺序交换后的新顺序仍然与NALU解码顺序一致，则它们可以具有相同的DON值。DON值相同的NALU可以以任意顺序解码，而DON值不同的两个NALU宜按上面规定的顺序被送给解码器。如果两个解码顺序连续的NALU具有不同的DON值，则第二个NALU的DON值宜等于第一个NALU的DON值加1。

注：接收端不宜期望两个连续的NALU（解码顺序）的DON值之差的绝对值为1，包括在没有传输错误的情况下。之所以不要求差值为1，是因为当确定NALU的DON值时，可能不知道是否所有的NALU都要发送到接收端。

4.6 单NALU包

这里定义的单NALU包必须只包含一个AVS-M视频标准中定义的NALU。复合包或分割单元不能在单NALU包中使用。通过按照RTP序列号顺序分解单NALU包得到的NAL单元流必须符合NALU解码顺序。单NALU包的结构如图2所示。

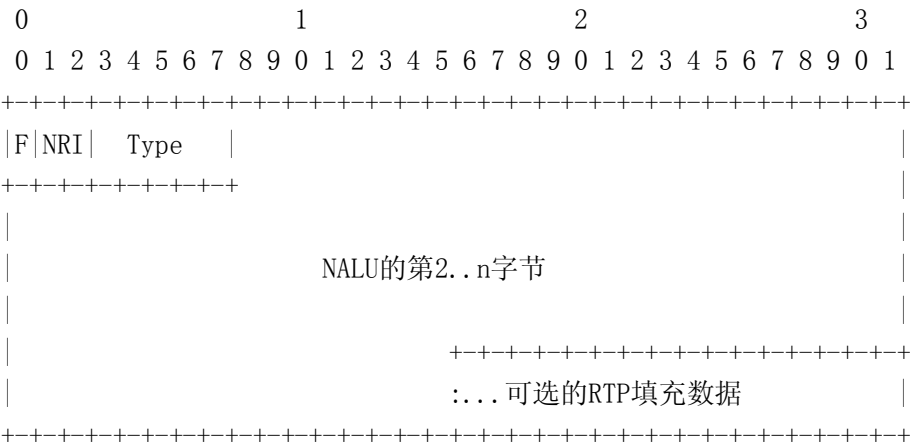


图2. 单NALU包的RTP净载格式

4.7 复合包

复合包的引入允许优化不同网络条件下的传输性能。例如，有线IP网络的MTU大小通常受限于以太网MTU大小，一般来说是1500字节，而有些无线通信系统的首选MTU大小为254字节或更小。引入复合包可以防止或减少不同网络系统之间的转码和不必要的打包头信息。

定义了以下两类复合包：

- 单时间复合包（STAP）：包含具有相同NALU时间的NALU。分为两个子类：STAP-A和STAP-B。STAP-A净载结构中不包含DON，而STAP-B净载结构中包含DON。
- 多时间复合包（MTAP）：包含可能具有不同NALU时间的NALU。根据NALU时间戳偏移的长度分为两个子类。

术语“NALU时间”定义为如果一个NALU单独打包为一个RTP包时的RTP时间戳。

复合包中的每个NALU被封装成一个复合单元。后面给出了四种不同的复合单元和它们的特征。

图3给出了复合包的RTP净载结构。

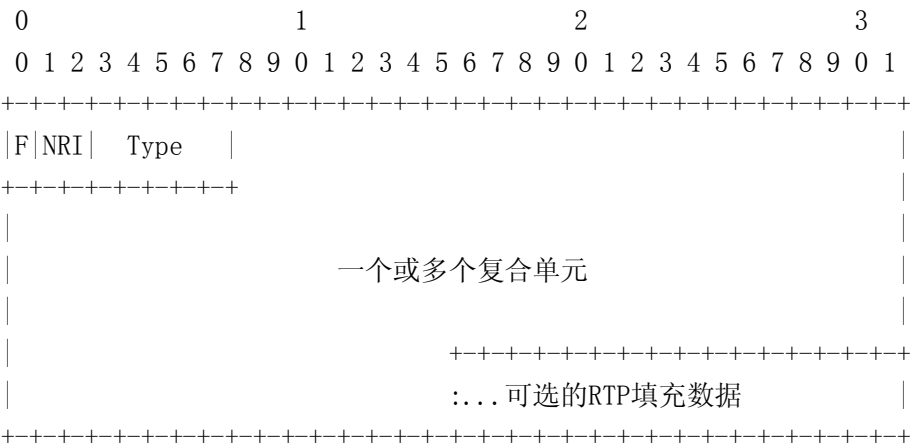






图5. STAP-B净载格式

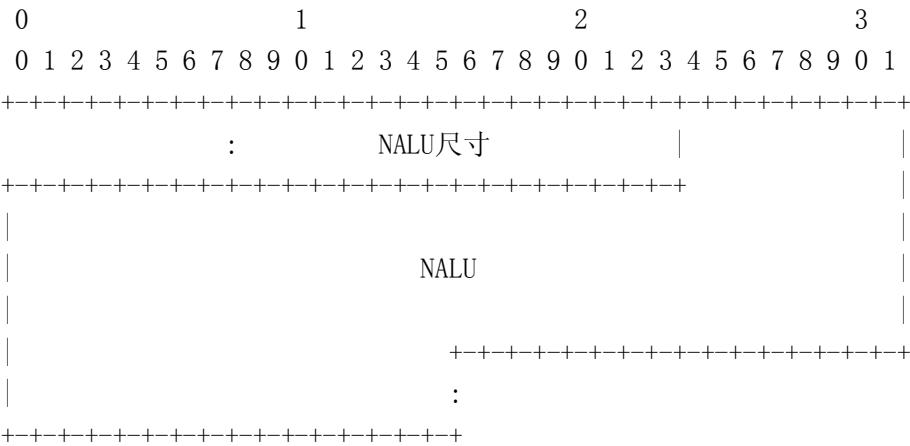


图6. 单时间复合单元的结构

图7示出了包含一个STAP-A的RTP包的例子。该STAP包含两个单时间复合单元，在图中标记为1和2。



图7. RTP包例子: 包含一个由两个单时间复合单元组成的STAP-A

图8示出了包含一个STAP-B的RTP包的例子。该STAP包含两个单时间复合单元，在图中标记为1和2。



图8. RTP包例子: 包含一个由两个单时间复合单元组成的STAP-B

4.7.2 多时间复合包

MTAP净载包含一个16位的无符号的解码顺序号基数（DONB，网络字节顺序）和至少一个多时间复合单元，如图9所示。DONB必须包含一个MTAP中第一个NALU（解码顺序）的DON值。

注：一个MTAP中按解码顺序的第一个NALU并不一定是按打包顺序的第一个NALU。

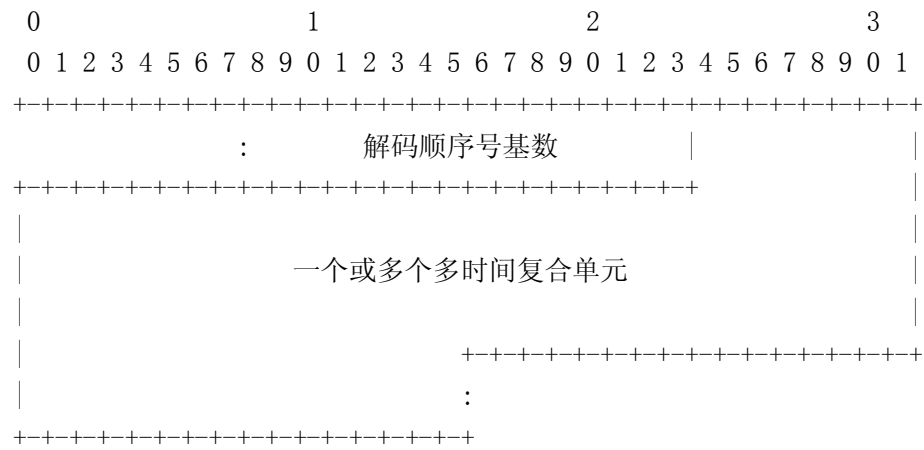


图9. MTAP净载格式

定义了两种多时间复合单元。一个复合单元包含一个16位（网络字节顺序）的无符号尺寸信息（描述随后的NALU的大小）、一个8位的无符号解码顺序号偏移（DOND）、和一个16或24位（网络字节顺序）的时间戳偏移。不同MTAP类型（MTAP16或MTAP24）的选择取决于具体应用，当时间戳偏移越大，MTAP的灵活性越高，但数据传输效率也会下降。

MTAP16和MTAP24的结构分别示于图10和图11。包中的复合单元的起始或结束位置不必是32位对齐的。复合单元中的NALU的DON值等于DONB与DOND值和模65536。本部分不规定MTAP中NALU的顺序，但在大多数情况下宜使用NALU解码顺序。

时间戳偏移的设定如下：如果NALU时间大于等于包的时间戳，则等于NALU时间减去RTP时间戳；否则等于NALU时间减去RTP时间戳再加上2<sup>32</sup>。

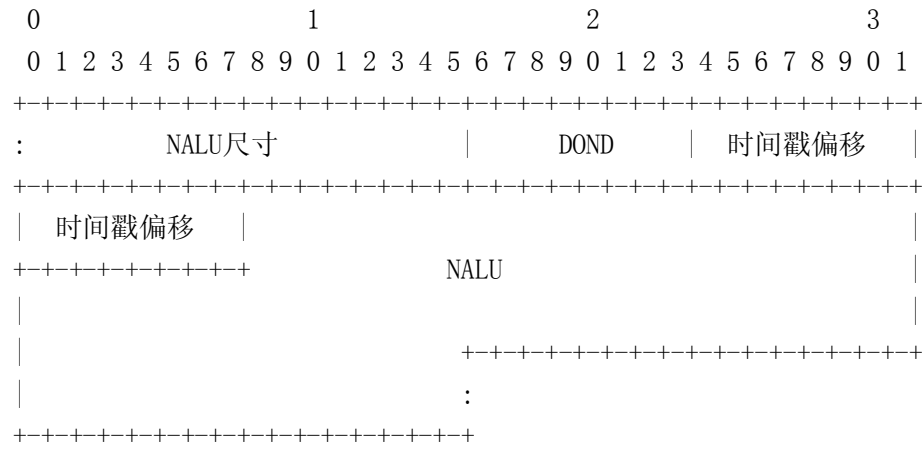


图10. 16位时间戳偏移的多时间复合单元





图13示出了包含一个MTAP24的RTP包的例子。该MTAP16包含两个多时间复合单元，在图中标记为1和2。



图13. RTP包：包含一个由两个多时间复合单元组成的MTAP24

## 4.8 分割单元

这个净载类型允许将一个NALU分割到多个RTP包中。在应用层进行分割而不是依赖于低层分割（如IP层），可以在IPv4网络中传输大于64K字节的NALU。大于64K字节的NALU可能存在于预先录制的视频流中，特别是在高分辨率格式下。

分割单元只适用于单个NALU而不适用于任何复合包。一个分割单元包含一个NALU的整数个字节，而NALU的每个字节必须完全包含在某一分割单元中。同一NALU的分割单元的发送必须以连续RTP序列号增长的顺序，即不能在这些分割单元中间插入发送其它RTP包。因此，一个NALU必须按照RTP序列号顺序组装。

一个被分割成分割单元的NALU叫做分割的NALU。STAP和MTAP不能分割。分割单元不能相互嵌套，即一个分割单元不能包含另一个分割单元。

包含一个分割单元的RTP包的RTP时间戳设为分割的NALU的NALU时间。

图14给出了FU-A的RTP净载格式。一个FU-A包含一个分割单元指示（1字节）、一个分割单元头（1字节）和一个分割单元净载。

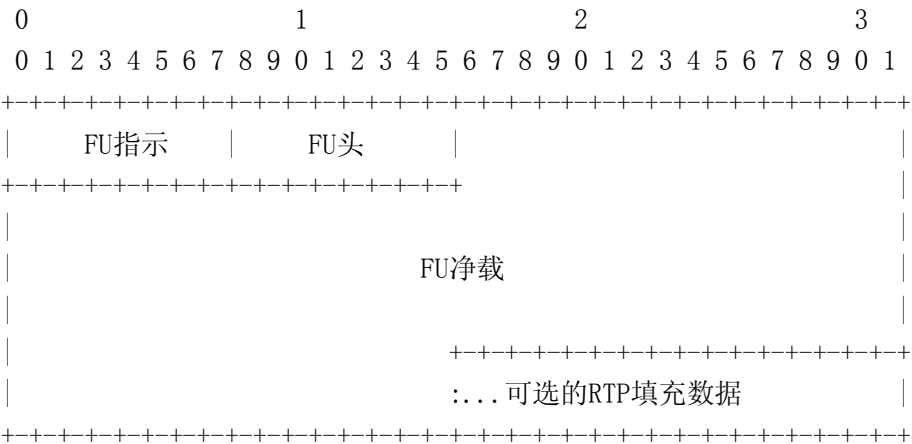


图14. FU-A的RTP净载格式

图15给出了FU-B的RTP净载格式。一个FU-B包含一个分割单元指示（1字节）、一个分割单元头（1字节）、一个解码顺序号（DON，网络字节顺序）和一个分割单元净载。

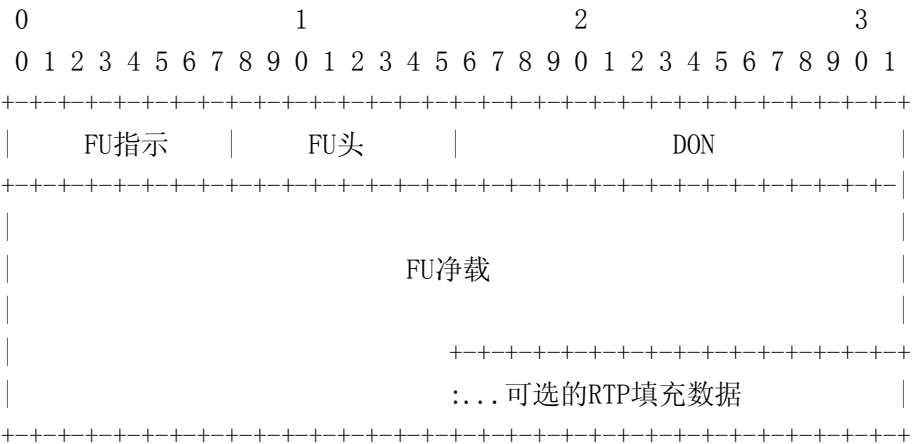
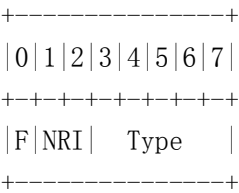


图15. FU-B的RTP净载格式

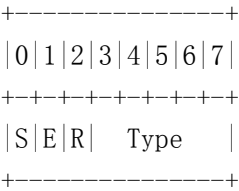
FU-B只能用于交织打包模式下作为一个分割的NALU的第一个分割单元，而不能用于任何其它场合。也就是说，在交织打包模式下，每个分割的NALU的第一个分割单元为FU-B，后跟一个或多个FU-A。

FU指示字节的格式为：



FU指示字节的Type域为28和29，分别识别FU-A和FU-B。F位的使用已在第4.3小节介绍。NRI域的设置与分割的NALU的NRI域相同。

FU头的格式如下：



**起始位 (S)：** 1位

为1表明一个分割的NALU的开始。如果随后的分割单元净载不是分割的NALU的开始，则设为0。

**结束位 (E)：** 1位

为1表明一个分割的NALU的结束，即净载的最后字节也是分割的NALU最后字节。如果随后的分割单元净载不是分割的NALU的最后分割单元，则设为0。

**保留位 (R)：** 1位

保留位必须为0，接收器应将其忽略。

**Type：** 5位

AVS-M视频标准的表7-1中所定义的NAL单元类型。

FU-B中的DON域的值的选择已在第4.5小节介绍。

注：FU-B中的DON域允许网关在没有将输入NALU组织成NALU解码顺序的情况下对NALU进行分割。

一个分割的NALU不能在同一个分割单元中传输，即同一分割单元头中的S位和E位不能同时为1。

将连续的分割单元的净载连接起来可以重建分割的NALU的净载。分割的NALU的头并不直接包含在分割单元净载中，而是包含在分割单元的FU指示字节的F和NRI域以及FU头的Type域中。一个分割单元净载可能包含任意数个字节，也可能是空的。

注：空的分割单元允许某类发送器在几乎无错的环境下减小延时。这类发送器的特点是它们在NALU完全产生之前（即知道NALU尺寸之前）打包分割单元。如果不允许空的分割单元，则发送器不得不在发送当前分割单元之前为随后的分割单元产生至少一个比特的数据。由于有时候一些宏块的编码比特数为0，这个限制会增加延时。不过，由于空的分割单元增加了传输NALU时的包的数目，从而增加了NALU数据丢失的几率，因此宜慎重使用。

如果一个分割单元被丢失，接收器宜丢弃随后（传输顺序）所有对应同一分割的NALU的分割单元。

终端或MANE中的接收器可以将一个NALU的前n-1个分割单元复合成一个不完全的NALU，即使该NALU的第n个分割单元没有收到。在这种情况下，NALU的forbidden\_zero\_bit必须设为1以表明语法错误。

5 打包规则

4.2小节介绍了打包模式。5.1小节介绍适用于一个以上打包模式的公共打包规则；5.2、5.3和5.4小节分别介绍适用于单NALU模式、非交织模式和交织模式的打包规则。

5.1 公共打包规则

无论使用那种打包模式，所有的发送器都必须遵守以下打包规则：

- 同一编码图像的图像头或条带NALU（这些NALU有相同的RTP时间戳值）可能以AVS-M视频标准相应档次所允许的任意顺序传送。然而，对于低延时系统，它们应按原始解码顺序传送以减小延时。
- 参数集按6.3小节给出的规则和推荐方法进行处理。
- 除了序列参数集、图像参数集或图像头NALU外，发送器（包括MANE）不能重复任何其它NALU。重复的序列和图像参数集不能影响任何激活的序列或图像参数集。重复的图像头后面必须紧跟该图像的条带NALU（不必为该图像的的第一个条带）。重复宜在应用层进行而不是通过重复RTP包（具有同样的序列号）。

采用非交织模式和交织模式的发送器必须遵守以下打包规则：

- MANE可以在RTP翻译器（translator）中将多个单NALU包转换成一个复合包、将一个复合包转换成多个单NALU包或将二者混合。RTP翻译器至少宜考虑如下参数：路径MTU大小、非均衡保护（unequal protection）机制（例如，根据IETF RFC 2733规定的基于包的FEC，尤其是用于参数集）、允许的系统延时和接收器的缓冲能力。

注：RTP翻译器必须按照IETF RFC 3550处理RTCP。

## 5.2 单 NALU 模式

当可选的MIME参数packetization-mode值等于0或packetization-mode不存在，并且没有外部方式确定使用其它打包模式时，使用单NALU打包模式。所有的接收器都必须支持这个模式。在这个模式下只能使用单NALU包，而不能使用STAP、MTAP和分割单元。单NALU包的传输顺序必须与NALU解码顺序一致。

## 5.3 非交织模式

当可选的MIME参数packetization-mode值等于1或有外部方式确定使用非交织模式时，使用非交织打包模式。这个模式宜被支持。在这个模式下只能使用单NALU包、STAP-A和FU-A，而不能使用STAP-B、MTAP和FU-B。NALU的传输顺序必须与NALU解码顺序一致。

## 5.4 交织模式

当可选的MIME参数packetization-mode值等于2或有外部方式确定使用交织模式时，使用交织打包模式。有些接收器可以支持这个模式。在这个模式下只能使用STAP-B、MTAP、FU-A和FU-B，而不能使用单NALU包和STAP-A。包和NALU的传输顺序已在4.5小节介绍。

## 6 净载格式参数

本节规定可以用来选择可选的净载格式特征和比特流的某些特征的参数。这些参数被规定为AVS-M视频标准的MIME子类型注册的一部分。为了便于某些应用采用SDP协议（Session Description Protocol, IETF RFC 2327），同时也给出了这些参数与SDP域之间的匹配。如果使用不涉及MIME或SDP的控制协议，则可以在其它地方定义等价的参数。

一些参数为接收器提供发送流的特性。这些参数的名字以“sprop”（即流特性，stream property）作为前缀。有些“sprop”参数受限于其它净载或解码配置参数。例如，参数sprop-parameter-sets受限于参数profile-level-id。所有“sprop”参数的选择由媒体发送器而不是接收器完成。“sprop”参数的这个特点可能不同于一些信令协议概念（在涉及这些信令协议的情况下宜避免使用这些参数）。

### 6.1 MIME 注册

AVS-M视频的MIME子类型从IETF树中分配。接收器必须忽略任何未规定的参数。

媒体类型名： video

媒体子类型名： AVSM

要求的参数： 无

可选的参数： profile-level-id, max-mbps, max-fs, max-cpb, max-dpb, max-br, sprop-parameter-sets, parameter-add, packetization-mode, sprop-interleaving-depth, sprop-deint-buf-req, deint-buf-cap, sprop-init-buf-time, sprop-max-don-diff, max-rcmd-nalu-size, Encoding considerations, Security considerations, Public specification, Additional information, File extensions, Macintosh file type code, Object identifier or OID, Person and email address to contact for further information, Intended usage, Author, Change controller.

#### 6.1.1 可选 MIME 参数的语义

##### profile-level-id:

AVS-M视频标准中规定的序列参数集中的下面两个字节的base16表示：profile\_idc和level\_idc。

如果profile-level-id用于表明NALU流的特性，则该参数表明为了解码时符合AVS-M视频标准一个解码器必须支持的档次和级别。

如果profile-level-id用于能力交换或会话建立过程，则该参数表明该编解码器支持的档次和支持的对应该档次的最高级别。

注：能力交换或会话建立过程宜提供分别为每个支持的档次列出能力的方法。例如，可以使用IETF RFC 3264第10.2小节介绍的编解码器选择方法。

如果profile-level-id不存在，缺省档次和级别为基本档次1.0级别。

#### **max-mbps, max-fs, max-cpb, max-dpb, and max-br:**

这些参数可以用来表明一个接收器的能力，而不能用作任何其它用途。包含这些参数中任何一个的能力描述必须同时包含profile-level-id参数。接收器必须完全支持profile-level-id表明的级别。如下所述，这五个参数可以用来表明接收器级别规定之外的扩展能力。

当这五个参数中多于一个参数存在时，接收器必须同时支持所有表明的能力。例如，如果max-mbps和max-br同时存在，则接收器支持由profile-level-id指定的相应的级别加上扩展的帧率和比特率。也就是说，接收器能够解码这样的NALU流，其最大宏块处理速度（即每秒最大宏块个数）为max-mbps，最大比特率为max-br，编码图像缓冲区大小按照下面max-br的语义规定的方法导出，而其它的特性符合相应的级别。

接收器给出的这五个参数不能同时符合下列条件：符合一个比由profile-level-id指定的相应的级别更高的级别的要求。

注：当可选的MIME参数用于表明NALU流的特性并且这五个参数不存在时，profile-level-id必须为这样的值：该NALU流完全符合指定的档次和级别。

#### **max-mbps:**

max-mbps的值是一个整数，指定每秒最大宏块个数，单位为宏块/秒。max-mbps表明接收器能够以高于profile-level-id指定的级别要求的速率进行解码。如果max-mbps存在，接收器必须能够解码这样的NALU流：除了AVS-M视频标准表B-3、B-4和B-5中的每秒最大宏块个数等于max-mbps以外，该NALU流符合相应的级别。max-mbps的值必须大于等于表B-3、B-4和B-5中给出的每秒最大宏块个数的值。发送器可能根据这个信息发送比相应级别指定的更高的帧率。

#### **max-fs:**

max-fs的值是一个整数，指定最大帧尺寸，单位为宏块。max-fs表明接收器能够解码大于profile-level-id指定的级别要求的图像尺寸。如果max-fs存在，接收器必须能够解码这样的NALU流：除了AVS-M视频标准表B-3、B-4和B-5中的每帧最大宏块个数等于max-fs以外，该NALU流符合相应的级别。max-fs的值必须大于等于表B-3、B-4和B-5中给出的每帧最大宏块个数的值。发送器可能根据这个信息发送比相应级别指定的尺寸更大的图像，此时帧率可能减小。

#### **max-cpb:**

max-cpb的值是一个整数，指定最大编码图像缓冲区大小，单位为1000比特。max-cpb表明接收器具有比profile-level-id指定的级别要求的最少编码图像缓冲区内更多的内存。如果max-cpb存在，接收器必须能够解码这样的NALU流：除了AVS-M视频标准表B-3、B-4和B-5中的最大编码图像缓冲区比特数等于1000\*(max-cpb)以外，该NALU流符合相应的级别。1000\*(max-cpb)的值必须大于等于表B-3、B-4和B-5中给出的最大编码图像缓冲区比特数的值。发送器可能根据这个信息生成比相应级别指定的最大编码图像缓冲区大小能够达到的更大的比特率变化范围。

注：编码图像缓冲区（CPB）用于AVS-M视频的虚拟参考解码器（HRD，AVS-M视频标准附录E）。HRD被推荐使用在编码器中以检验产生的比特流符合标准和控制输出比特率。因此，概念上CPB独立于发送器中其它可能的缓冲区，包括解交织（de-interleaving）缓冲区和解抖动（de-jitter）缓冲区。CPB在解码器上的实现不必按照AVS-M视频标准附录E的规定，只要能够解码符合标准的比特流，符合标准的解码器可以使用任何缓冲方法。因此，在实践过程中，视频解码器的输入缓冲区可以与接收器的解交织缓冲区和解抖动缓冲区结合在一起。

**max-dpb:**

max-dpb的值是一个整数，指定最大解码图像缓冲区大小，单位为1024字节。max-dpb表明接收器具有比profile-level-id指定的级别要求的最少解码图像缓冲区内更多的内存。如果max-dpb存在，接收器必须能够解码这样的NALU流：除了AVS-M视频标准表B-3、B-4和B-5中的最大解码图像缓冲区字节数等于 $1024 * (\text{max-dpb})$ 以外，该NALU流符合相应的级别。 $1024 * (\text{max-dpb})$ 的值必须大于等于表B-3、B-4和B-5中给出的最大解码图像缓冲区字节数的值。发送器可能根据这个信息生成比相应级别指定的最大解码图像缓冲区大小能够达到的更高的压缩效率。

**max-br:**

max-br的值是一个整数，指定最大视频比特率，单位为1000比特。max-br表明接收器的视频解码器能够解码比profile-level-id指定的级别要求的更高的比特率。如果max-br存在，接收器的视频解码器必须能够解码这样的NALU流：该NALU流符合相应的级别，除了

- AVS-M视频标准表B-3、B-4和B-5中的最大比特率等于 $1000 * (\text{max-br})$ ；
- 如果max-cpb不存在，AVS-M视频标准表B-3、B-4和B-5中的最大编码图像缓冲区比特数等于原值乘以 $(1000 * (\text{max-br}) / (\text{相应级别规定的最大比特率}))$ ，其中相应级别规定的最大比特率的单位为比特/秒。

$1000 * (\text{max-br})$ 的值必须大于等于表B-3、B-4和B-5中给出的最大比特率的值。发送器可能根据这个信息发送比相应级别指定的更大的最大比特率，从而提高视频质量。

**sprop-parameter-sets:**

这个参数可以用来传输任何初始序列和图像参数集NALU，这些NALU在任何其它NALU之前（解码顺序）。该参数不能用在任何能力交换过程中表明编解码器的能力。该参数的值为初始参数集NALU的base64表示。这些参数集按解码顺序排列，每两个参数集中间插入一个半角逗号','。

**parameter-add:**

这个参数可以用来表明是否允许这个参数的接收者在它的信令响应中使用sprop-parameter-sets添加参数集。值为0或1。0表示不允许添加参数集；1表示允许添加参数集。如果该参数不存在，其缺省值为1。

**packetization-mode:**

这个参数可以用来表明一个RTP净载类型的特性或接收器的能力。由于只能指出一个配置点，当表明具有支持多种打包模式的能力时，必须使用多个配置点（多个RTP净载类型）。

值为0或不存在时，必须使用第5.2小节介绍的单NALU模式；值为1时必须使用第5.3小节介绍的非交织模式；值为2时必须使用第5.4小节介绍的交织模式。packetization-mode的取值范围为0..2。

**sprop-interleaving-depth:**

当packetization-mode不存在或packetization-mode值为0或1时，sprop-interleaving-depth必须不存在。当packetization-mode值为2时，sprop-interleaving-depth必须存在。

sprop-interleaving-depth表明NALU流的特性。它指定相对于任一NALU传输顺序在前而解码顺序在后的NALU数目的最大值。因此，可以保证接收器能够采用大于等于sprop-interleaving-depth+1个NALU的解码顺序恢复缓冲区来重建NALU解码顺序。

sprop-interleaving-depth的取值范围为0..32767。

**sprop-deint-buf-req:**

当packetization-mode不存在或packetization-mode值为0或1时，sprop-deint-buf-req必须不存在。当packetization-mode值为2时，sprop-deint-buf-req必须存在。

sprop-deint-buf-req表明NALU流需要的解交织缓冲区尺寸，单位为字节。该参数的值必须大于等于10.2 小节描述的解交织缓冲区需要的最大缓冲区充满度（单位为字节）。

sprop-deint-buf-req的取值范围为0..4294967295。

注：sprop-deint-buf-req仅表明需要的解交织缓冲区的大小。如果发生网络抖动，还需要一个适当大小的抖动缓冲区。

#### **deint-buf-cap:**

这个参数表明接收器的能力,指定接收器具有的用来重建NALU解码顺序的解交织缓冲区空间的大小（单位为字节）,并表明接收器能够处理任何sprop-deint-buf-req值小于等于deint-buf-cap的流。

如果参数不存在，则deint-buf-cap的缺省值为0。deint-buf-cap的取值范围为0..4294967295。

注：deint-buf-cap仅表明接收器可以提供的解交织缓冲区的尺寸的最大值。如果发生网络抖动，还需要一个适当大小的抖动缓冲区。

#### **sprop-init-buf-time:**

这个参数可以用来表明NALU流的特性。如果packetization-mode值为0或1，sprop-init-buf-time必须不存在。

该参数表明接收器开始解码以从传输顺序恢复NALU解码顺序前所必须经过的初始缓冲时间。假设传输是可靠的和瞬时的、传输和解码的时间线（timeline）相同以及第一个包到达时解码开始，sprop-init-buf-time等于一个NALU的传输时间和解码时间之差的最大值。

下面给出了一个指定sprop-init-buf-time值的例子。NALU流的传输顺序为从左到右，数值表示解码时间：

0 2 1 3 5 4 6 8 7 ...

假设NALU传输速率恒定，则传输时间为：

0 1 2 3 4 5 6 7 8 ...

让传输时间和解码时间按列相减，得到：

0 -1 1 0 -1 1 0 -1 1 ...

因此，这个例子中sprop-init-buf-time的值为1，单位为NALU传输时间间隔。

sprop-init-buf-time为非负整数，单位为1/90000秒，取值范围为0..4294967295。如果该参数不存在，则初始缓冲时间未定义。

除了sprop-init-buf-time，接收器宜考虑各种网络单元产生的传输延时抖动。

#### **sprop-max-don-diff:**

这个参数可以用来表明NALU流的特性，但不能用来表明发送器、接收器或编解码器的能力。如果packetization-mode值为0或1，sprop-init-buf-time必须不存在。sprop-max-don-diff的取值范围为0..32767。该参数不存在表示参数值未规定。sprop-max-don-diff的计算方法如下：

$$\text{sprop-max-don-diff} = \max\{\text{AbsDON}(i) - \text{AbsDON}(j)\}, \text{ 任意 } i \text{ 和任意 } j > i$$

其中i和j表示NALU的传输顺序索引，AbsDON表示NALU的绝对解码顺序号，即其值到达65535后继续增加而不重置为0。AbsDON的计算方法如下：

令m和n为传输顺序连续的两个NALU。对于传输顺序上的第一个NALU（其索引为0）， $\text{AbsDON}(0) = \text{DON}(0)$ ；对于其它的NALU，

If  $\text{DON}(m) == \text{DON}(n)$ ,  $\text{AbsDON}(n) = \text{AbsDON}(m)$

If  $(\text{DON}(m) < \text{DON}(n) \text{ and } \text{DON}(n) - \text{DON}(m) < 32768)$ ,  
 $\text{AbsDON}(n) = \text{AbsDON}(m) + \text{DON}(n) - \text{DON}(m)$

If  $(\text{DON}(m) > \text{DON}(n) \text{ and } \text{DON}(m) - \text{DON}(n) \geq 32768)$ ,  
 $\text{AbsDON}(n) = \text{AbsDON}(m) + 65536 - \text{DON}(m) + \text{DON}(n)$

If  $(\text{DON}(m) < \text{DON}(n) \text{ and } \text{DON}(n) - \text{DON}(m) \geq 32768)$ ,  
 $\text{AbsDON}(n) = \text{AbsDON}(m) - (\text{DON}(m) + 65536 - \text{DON}(n))$

If ( $DON(m) > DON(n)$  and  $DON(m) - DON(n) < 32768$ ),  
 $AbsDON(n) = AbsDON(m) - (DON(m) - DON(n))$

其中DON(i)传输顺序索引为i的NALU的解码顺序号，见4.5小节。

注：接收器可能利用sprop-max-don-diff推知缓冲区中哪些NALU可以传递给解码器。

**max-rcmd-nalu-size:**

这个参数可以用来表明接收器的能力，而不能用作任何其它目的。该参数值指定接收器能够有效处理的最大NALU尺寸，单位为字节。该参数值是个推荐值而并不是一个严格的上限。发送器可以产生更大的NALU但是必须知道其处理的代价可能更高。

max-rcmd-nalu-size必须为整数，取值范围为0..4294967295。如果该参数不存在，则其值为未知。发送器仍然必须考虑发送器和接收器之间的MTU尺寸。为此，发送器宜运行MTU检测手段。

注：将此参数设为比需要更低的值可能带来负面影响。

**Encoding considerations:**

这个类型的定义仅仅为了RTP (IETF RFC 3550) 传输。

**Security considerations:**

见本部分第7节。

**Public specification:**

请参考本部分和其第2节。

**Additional information:**

无

**File extensions:**

无

**Macintosh file type code:**

无

**Object identifier or OID:**

无

**Person & email address to contact for further information:**

[ye-kui.wang@nokia.com](mailto:ye-kui.wang@nokia.com)

**Intended usage:**

普通 (COMMON)

**Author:**

[ye-kui.wang@nokia.com](mailto:ye-kui.wang@nokia.com)

**Change controller:**

AVS System Subgroup

## 6.2 SDP 参数

### 6.2.1 MIME 参数与 SDP 域之间的匹配

MIME媒体类型video/AVSM字符串与SDP (Session Description Protocol, 会话描述协议) 域之间的匹配方法为：

- SDP "m=" 行的媒体名必须为video。
- SDP "a=rtpmap" 行的编码名必须为AVSM (MIME子类型)。
- SDP "a=rtpmap" 行的时钟率 (clock rate) 必须为90000。



- 如果存在，可选参数profile-level-id、max-mbps、max-fs、max-cpb、max-dpb、max-br、sprop-parameter-sets、parameter-add、packetization-mode、sprop-interleaving-depth、deint-buf-cap、sprop-deint-buf-req、sprop-init-buf-time、sprop-max-don-diff和max-rcmd-nalu-size必须包括在SDP "a=fmtp" 行。这些参数被表示为MIME媒体类型字符串，表示形式为一系列由半角分号';'分开的“参数=值”对。

下面给出了一个SDP媒体描述的例子（基本档次级别1.0）：

```
m=video 49170 RTP/AVP 98
a=rtpmap:98 AVSM/90000
a=fmtp:98 profile-level-id=10A0;
      sprop-parameter-sets=[SPS#0],[PPS#0]
```

其中“[SPS#0],[PPS#0]”表示由半角逗号分开的一个序列参数集和一个图像参数集的base64表示。

### 6.2.2 在 SDP 提供/回答模型中的用法

在单点传输（unicast）情形下，如果AVS-M视频通过RTP提供并且在协商（negotiation）过程中使用SDP提供/回答（Offer/Answer）模型，则必须遵守以下规则：

- 识别 AVS-M 视频媒体格式配置 的参数为 profile-level-id、packetization-mode 和 sprop-deint-buf-req（如果packetization-mode等于2）。这三个参数必须对称地使用：即回答者必须保留所有的配置参数；或完全去除该媒体格式（净载类型），如果不支持一个或多个参数值。

注：对称使用的要求仅适用于上述三个参数，而不适用于其它流特性和能力参数。

为了简化这些配置的处理和匹配，提供中使用的RTP净载类型值宜同样在回答中使用，如IETF RFC 3264中所规定。除非由上述三个参数决定的配置相同，回答中不能包含提供中用到的净载类型值。

注：收到回答后，提供者必须根据媒体类型（即video/AVSM）和上述三个参数比较没有包含在提供中的净载类型与已经包括在提供中的净载类型，以确定相应配置是新的还是与已提供的配置相同。

- sprop-parameter-sets、sprop-deint-buf-req、sprop-interleaving-depth、sprop-max-don-diff和sprop-init-buf-time表明提供者或回答者发送的具有某一配置的NALU流的特性。这有别于提供/回答参数的一般用途——表明提供者或回答者能够接收的流的特性。对于AVS-M视频，提供者假定回答者能够接收采用所提供配置编码的媒体。

注：上面的参数适用于申明这些参数的实体发送的任何具有相同配置的流。即这些参数并不受限于净载类型，而是可能适用于来自同一实体的具有相同配置的其它净载类型。

- 能力参数max-mbps、max-fs、max-cpb、max-dpb、max-br和max-rcmd-nalu-size可以用来表明其它能力。对它们的解释取决于方向属性。如果方向属性为sendonly，则这些参数描述发送器能够产生的RTP包和NALU流的限制。如果方向属性为sendrecv或recvonly，则这些参数描述接收器能够接受的限制。
- 如前所述，提供者必须在提供中包含交织的AVS-M视频流的解交织缓冲区的尺寸。为了使提供者和回答者都能够告诉对方自己在解交织缓冲方面的能力，推荐二者都包含deint-buf-cap。这个信息可以在第二轮提供和回答过程中选择了sprop-deint-buf-req的值之后使用。对于交织流，如果接收器的能力未知，推荐考虑提供多个缓冲需求不同的净载类型。
- 参数sprop-parameter-sets的使用如前所述。另外，回答者必须在它的回答中保留所有从提供中收到的参数集。如果parameter-add为0，回答中不能增加任何参数集；如果parameter-add为1，回答者可以在它的回答的sprop-parameter-sets中增加参数集。无论parameter-add的值为0或1，回答者必须答应接收采用包含在自己的回答中的sprop-parameter-sets的视频流。

注：增加参数集时必须小心，不能由于使用了冲突的参数集标号而重写已经传输的参数集。

对于通过多点传送（multicast）方式发布的流，必须额外遵守以下规则：

- 回答者不能改变流特性参数 sprop-parameter-sets、sprop-deint-buf-req、sprop-interleaving-depth、sprop-max-don-diff和sprop-init-buf-time。因此，一个提供的净载类型要么被原封不动地接受，要么被去除。
- 回答者必须对所有申明为sendrecv或recvonly的流支持接收器能力参数max-mbps、max-fs、max-cpb、max-dpb、max-br和max-rcmd-nalu-size。否则，回答者必须去除该媒体格式或拒绝会话。

下面完整地列出了参数在不同情形下的意义。

- 在“a=sendrecv”或未使用方向属性或“a=recvonly”的提供和回答中各参数的意义如下：

申明要接收的配置或特性：

- profile-level-id
- packetization-mode

申明要发送的流的特性（仅当“a=sendrecv”或未使用方向属性时适用）：

- sprop-deint-buf-req
- sprop-interleaving-depth
- sprop-parameter-sets
- sprop-max-don-diff
- sprop-init-buf-time

申明接收器实现的能力：

- max-mbps
- max-fs
- max-cpb
- max-dpb
- max-br
- deint-buf-cap
- max-rcmd-nalu-size

申明如何进行提供/回答协商：

- parameter-add

- 在“a=sendonly”的提供和回答中各参数的意义如下：

申明提议要发送的流的配置和属性：

- profile-level-id
- packetization-mode
- sprop-deint-buf-reg
- sprop-max-don-diff
- sprop-init-buf-time
- sprop-parameter-sets
- sprop-interleaving-depth

申明发送器接收流时的能力：

- max-mbps
- max-fs
- max-cpb
- max-dpb
- max-br
- deint-buf-cap
- max-rcmd-nalu-size

申明如何进行提供/回答协商：

- parameter-add

另外，还需要考虑下面的因素。

- 申明接收器能力的参数通常是可以降低等级的，即它们表达发送器可能的行为的上限。因此发送器可以选择设定它的编码器只使用这些参数的更小(低)或相等的值。sprop-parameter-sets不能用来申明发送器的能力。
- 除了profile-level-id的级别(level)部分以外，申明配置点的参数不可以降低等级。这个表明了接收器期望使用的值，发送器端必须原封不动地使用这些值。
- 如果申明了发送器的能力，并且申明中使用了不可降低等级的参数，则这些参数表明了一个可接受的配置。为了达到更好的互用性(interoperability)，可以提供多个可替换的配置，例如不同的打包模式。由于不可能采用同一个净载类型提供多个配置，如果提供多个配置，每个配置需要自己的RTP净载格式。
- 接收器宜理解一个净载格式所有的MIME参数，即使它只支持该净载格式的一部分功能。这保证了接收器能够理解一个提供的流可以被降低等级到自己能够支持的程度。
- 回答者可以扩展额外的媒体格式配置到提供中。然而，为了能够使用这些额外的配置，大多数情形下需要提供者的第二个提供，以给出媒体发送器要使用的流特性参数。这个过程还有另外一个效果，就是提供者必须能够接收这个媒体格式配置，而不仅仅是发送。
- 如果提供者希望发送和接收的能力不对称，则它必须提供不同的RTP会话，即分别申明为recvonly和sendonly的不同SDP “m=”行。

### 6.2.3 在申明性的会话描述中的用法

如果AVS-M视频RTP流的SDP提供方式为申明性的，例如在RTSP和SAP中，则需要考虑下面的因素。

- 所有能够表明NALU流和接收器特性的参数都用来表明NALU流的特性。例如，在这种情况下，参数profile-level-id申明流使用的值，而不是发送器的能力。因此，参数的意义如下：

申明配置或特性：

- profile-level-id
- sprop-parameter-sets
- packetization-mode
- sprop-interleaving-depth
- sprop-deint-buf-reg
- sprop-max-don-diff
- sprop-init-buf-time

不可用的：

- max-mbps
- max-fs
- max-cpb
- max-dpb
- max-br
- max-rcmd-nalu-size
- parameter-add
- deint-buf-cap

- SDP接收器需要支持提供的所有参数和参数值。否则接收器必须拒绝(RTSP)或不参加(SAP)该会话。采用期望接收应用支持的值是会话创建者的责任。

### 6.3 参数集考虑

参数集对AVS-M视频编解码器的操作起关键作用。由于其特点和对解码过程的重要性，参数集的丢失或错误传输造成的影响很难在接收端进行掩蔽。引用一个错误的参数集通常会对解码过程造成非常严重的结果。这种错误可能是由于错误传输或丢失一个参数集数据，也可能是由于参数集更新的非及时传输。因此，下面为RTP发送器的实现推荐了一些指导原则。

参数集NALU的传输可以通过三种不同的原则：

- A. 采用会话控制协议在实际RTP会话之前带外（out-of-band）传输。
- B. 采用会话控制协议在实际RTP会话过程中带外（out-of-band）传输。
- C. 在实际RTP会话过程中作为RTP流的一部分带内（in-band）传输。

A和B原则要在会话控制协议中实现。C原则由本部分定义的RTP净载格式支持。

序列和图像参数集NALU不宜在RTP净载中传输，除非为RTP提供了可靠的传输，因为丢失一个参数集很可能导致RTP流的相当多的数据的解码失败。因此，推荐使用基于可靠的会话控制协议传输参数集的方法，即A和B原则。

本节的后续部分假定带外参数集传输可靠而带内传输不可靠。如果采用了带内传输，发送者宜考虑错误特性采用一些方法提高正确传输参数集的概率。这些方法包括包重复、前向纠错码和重传。采用不可靠的带外控制协议具有与带内传输相似的缺点，而且还可能导致同步上的困难（见下面），因此不推荐使用。

会话过程中可能采用B或C原则添加或更新参数集。参数集需要被引用之前到达解码器。为了避免更新或添加参数集可能导致的更多的问题，宜考虑以下推荐措施：

- 添加或更新参数集时，C原则对传输错误敏感，因此推荐使用B原则。
- 添加或更新参数集时，宜注意保证任何参数集在被引用前到达。通常带外和带内信号之间不存在同步。如果采用带外传输，推荐在收到参数集到达的确认信号之前不开始传送需要相应参数集的NALU。
- 参数集更新之后，宜考虑以下同步问题。在接收端重写一个参数集之前，发送器必须保证网络或接收器缓冲区中的任何NALU不再需要被重写的参数集。否则可能发生利用错误的参数集解码的情况。为了解决这个问题，推荐只重写那些足够长时间没有被引用的参数集（以保证所有相关的NALU已经解码），或添加一个新的参数集（这可能对视频编码效率造成负面影响）。
- 添加参数集时利用以前未用的参数集标号。这样可以避免上面提到的问题。然而，在多点会话中，除非采用了同步的控制协议，可能出现多个实体试图采用同一标号添加不同的参数集的问题，这个必须避免。
- 同时采用B和C原则在同一RTP会话中添加或更新参数集可能导致参数集的不一致，因为控制信道和RTP信道之间没有同步。因此，B和C原则不能同时使用于同一会话中，除非提供了足够的同步。

在有些情形下，使用带外参数集传输是不可能的，即参数集必须带内传输。这时参数集与其它数据的同步是固有的隐含的，但是必须考虑丢失的可能性。宜采用前面提到的方法减小丢失的概率。

如果初始参数集采用A原则提供，而后来采用带内（C原则）方法添加或更新，则更新带外传输的参数集时存在以下危险。如果接收器没有收到一些带内更新的参数集（例如由于丢失或较晚加入会话），这些接收器将试图利用过时的参数集解码。因此，推荐参数集标号在带外传输和带内传输的参数集之间分开使用。

为了使AVS-M视频编码器可以达到最高的灵活性和性能，如果可能，推荐允许任何发送器在会话中添加和使用自己的参数集。除非在会话拓扑结构不允许会话参与者添加自己的参数集，否则不宜将参数parameter-add设为0。

## 7 安全考虑

采用本部分定义的净载格式的RTP包遵守IETF RFC 3550以及任何适当的RTP档次，如IETF RFC 3551，中讨论的安全考虑。这意味着媒体流的保密性通过加密得到，例如采用IETF RFC 3711。因为本净载格式对数据压缩的采用是端到端的，加密需要在压缩之后进行。

一个可能的攻击方式采用导致接收端计算量不一致的压缩技术的数据编码。攻击者可以向流中注入病态的数据，这些数据的解码复杂，可以导致接收器超负荷。AVS-M对这种攻击尤其敏感，因为很容易产生一些NALU以影响很多其它NALU的解码过程。因此，推荐至少对RTP包进行数据源鉴别和数据完整性保护，例如采用IETF RFC 3711。

注意保证RTP包和它们的净载的保密性和完整性的适当机制取决于具体应用和相应的传输与信令协议。因此，尽管上面的例子中给出了IETF RFC 3711，其它的选择也是可能的。

解码器必须谨慎处理用户数据SEI消息，尤其当它们包含活跃的元素时。必须限制它们对流播放过程的适用范围。

端到端的安全如鉴定、完整性或保密性保护使得MANE除丢弃完整的包以外不能进行其它基于媒体知识的操作。在保护保密性的情形下，甚至基于媒体知识丢弃包也不能进行。为了允许任何MANE进行操作，要求允许该MANE是包含在安全环境建立过程中的可以信任的实体。

## 8 拥塞控制

RTP拥塞控制的采用必须根据IETF RFC 3550以及任何适当的RTP档次，如IETF RFC 3551。如果采用尽力服务（best-effort service）模式，一个额外的需求是：净载格式的使用者必须监视包丢失状况，以保证丢包率在可接受的参数范围内。在下面这种情况下认为丢包率可以接受：如果通过同一网络路径和相同网络条件的TCP流在合理的时间范围内达到不低于RTP流达到的平均吞吐量。这个条件可以通过以下方式满足：实现调节传输率或分层多播会话订购层数的拥塞控制机制，或当丢包率过高时安排一个接收者离开会话。

遵守拥塞控制需要的比特率调节在实时编码情况下很容易实现。然而，在传输预先编码的内容时，带宽调节需要存在同一内容的不同比特率的编码描述，或比特流中存在非参考图像。不同描述之间的切换通常可以在同一RTP会话中进行，如在IDR或GDR随机访问点进行。仅当不可降低等级的参数，如档次级别中的档次，需要改变时，才需要结束再重新开始媒体流。这个可以通过使用不同的RTP净载类型实现。

## 9 IANA 注册

将请求IANA注册一个新的MIME类型，见6.1小节。

## 10 解包过程（资料性）

解包过程由具体实现决定。因此，下面描述的宜视为一个合适的实现。其它方法也可以采用。对所介绍的算法进行优化也是可能的。10.1小节介绍单NALU和非交织打包模式的解包过程；10.2小节给出交织打包模式的解包过程；10.3小节介绍了额外的接收器解包指导原则。

所有与缓冲区管理有关的常用RTP机制适用于此。尤其是，重复的和过时的RTP包（根据RTP序列号和时间戳）将被丢弃。为了确定确切的解码时间，必须考虑为了适当的媒体流之间的同步而故意允许的延时等因素。

### 10.1 单 NALU 和非交织模式

接收器包含一个接收缓冲区以补偿传输延时抖动。接收器按接收顺序将包存入接收缓冲区。解包按RTP序列号顺序进行。对于单NALU包，解包得到的NALU直接传递给解码器。对于STAP-A，解包得到的NALU按它们在包中出现顺序传递给解码器。对于FU-A，解包得到的分割单元与以前解包得到的同一NALU的分割单元（如果存在）连接，如果连结的结果为完整的NALU则传递给解码器。

## 10.2 交织模式

这些解包规则的概念实际上就是将NALU从传输顺序重排序到NALU解码顺序。

接收器包含一个接收缓冲区以补偿传输延时抖动和将NALU从传输顺序重排序到NALU解码顺序。本节在假设没有传输延时抖动的情况下描述接收器操作。为了区别于实际中同时用于补偿传输延时抖动的接收缓冲区，这儿称之为解交织缓冲区。

### 10.2.1 解交织缓冲区尺寸

如果在会话建立时采用了SDP提供/回答模型或任何其它能力交换过程，接收到的流的特性不宜超过接收器的能力。在SDP提供/回答模型中，接收器可以通过MIME参数deint-buf-cap表明其分配解交织缓冲区的能力。发送器通过MIME参数sprop-deint-buf-req表明对解交织缓冲区尺寸的需求。因此，推荐将解交织缓冲区的尺寸（单位为字节）设为大于等于sprop-deint-buf-req的值。

如果在会话建立时采用了申明性的会话描述，MIME参数sprop-deint-buf-req仍然表明对解交织缓冲区尺寸的需求。因此，同样推荐将解交织缓冲区的尺寸（单位为字节）设为大于等于sprop-deint-buf-req的值。

### 10.2.2 解交织过程

接收器有两种缓冲状态：初始缓冲和回放缓冲。初始缓冲发生于RTP会话初始化时。初始缓冲后，解码和回放开始，此时的缓冲状态为回放缓冲。

无论缓冲状态如何，接收器按接收顺序将包存入解交织缓冲区的方法如下：复合包中的NALU逐个存入解交织缓冲区；每个NALU的DON值被计算出来并存起来。

借助于下面的函数和常数，接收器的操作描述如下：

- 函数AbsDON的定义见6.1小节。
- 函数don\_diff的定义见4.5小节。
- 常数N等于MIME参数sprop-deint-buf-req（见6.1小节）的值加1。

直到下列条件之一成立，初始缓冲结束：

- 解交织缓冲区中有N个NALU。
- 如果sprop-max-don-diff存在， $\text{don\_diff}(m, n)$  大于sprop-max-don-diff的值，其中n表示接收的NALU中AbsDON值最大的NALU，而m表示接收的NALU中AbsDON值最小的NALU。
- 初始缓冲的持续时间等于或超过sprop-init-buf-time的值。

确定要从解交织缓冲区中移出的NALU的方法为：

- 如果解交织缓冲区中有至少N个NALU，则按下面的顺序移出NALU直到缓冲区中只剩下N-1个NALU。
- 如果sprop-max-don-diff存在，所有使得 $\text{don\_diff}(m, n)$  大于sprop-max-don-diff的NALU按下面的顺序移出解交织缓冲区并传递给解码器，其中m表示待测NALU，n表示解交织缓冲区中AbsDON值最大的NALU。

上面提到的将NALU从解交织缓冲区中移出并传递给解码器的顺序为：

- 令变量PDON在RTP会话开始时初始化为0。
- 对每个具有DON值的NALU计算DON距离如下：如果该NALU的DON值大于PDON值，DON距离等于(DON - PDON)；否则，DON距离等于(65535 - PDON + DON + 1)。
- NALU按DON距离增加的顺序从解交织缓冲区中移出并传递给解码器。如果几个NALU具有相同的DON距离，它们可以以任何顺序从解交织缓冲区中移出和并传递给解码器。

- 当解交织缓冲区中只剩下N-1个NALU时，PDON值设为最后一个从解交织缓冲区中移出和并传递给解码器的NALU的DON值。

### 10.3 额外的解包指导原则

可以采用下面的额外解包指导原则来实现实际的AVS-M视频解包器：

- RTP接收器（如网关）可能检测丢失的分割单元。如果发现了丢失的分割单元，网关可以决定不传送同一分割的NALU的后续分割单元，因为它们对AVS-M视频解码器已经没有意义。这样，MANE可以无需解析复杂的比特流丢弃无用包，从而减小网络负载。
- 必须丢弃包或NALU的接收器宜首先丢弃所有NALU头中NRI值为0的包或NALU。这样将减小对用户体验的影响并保留参考图像。如果需要丢弃更多的包，宜先丢弃NRI值较小的包，然后再丢弃NRI值较大的包。然而，丢弃任何NRI值大于0的包都可能引起解码错误传播，因此宜避免。