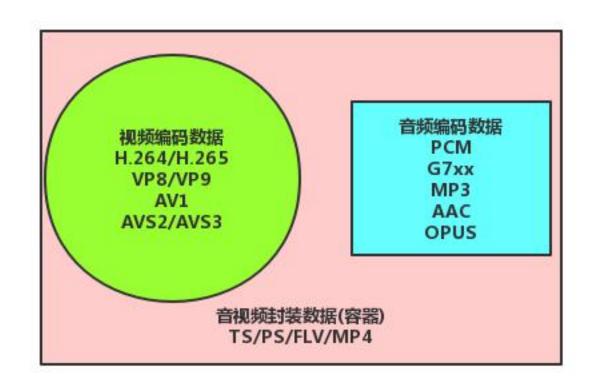
# 音视频封装面面观

X-Lab:田野 2019/12/17 概要

•音视频基础

•音视频封装格式

## 封装基础





## 封装基础

- 0.该封装格式推出机构和Spe
- 1.该封装格式支持的音视频编码格式以及怎么封装的
- 2.封装格式的复杂性、容错性, 灵活性
- 3.封装格式对流化和网络传输协议的支持
- 4.封装格式对码流高标清切换,对Trick-play播放的支持
- 5.封装格式里面时间戳信息是怎么体现的
- 6.不同封装格式对使用场景的支持

## 封装总结

| 名称  | 推出机构 时间                          | 支持视频编码格式  | 支持音频编<br>码格式        | 传输流媒体协议                         | 概括性描述   | 目前使用领域                 |
|---|----------------------------------|---|---------------------|---------------------------------|---|------------------------|
| TS<br>MPEG2-TS<br>Transport Stream          | MPEG组织<br>1994年                  | MPEG2 MPEG4H. 263<br>H. 264                         | G7xx<br>AAC         | HLS (TCP)                       | VCD/DVD时代的封装格式,逐<br>渐推出历史舞台,得益于苹果<br>还能存活一段时间。              | 广电<br>数字电视<br>苹果系列产品安防 |
| PS<br>MPEG2-PS<br>Program Stream            | MPEG组织<br>1994年                  | MPEG2 MPEG4H. 263<br>H. 264                         | G7xx<br>AAC         | RTP (TCP UDP)<br>RTSP           | VCD/DVD时代的封装格式,基本已经退出历史舞台,安防还有少许使用。                         | DVD电影<br>安防            |
| FLV<br>Flash Video                          | Adobe                            | Н. 264  | MP3<br>AAC          | HTTP (TCP)<br>RTMP (TCP)        | Adobe目前已经不再更新标准,<br>CDN存量比较大,还有一定的<br>市场,但是长远看也会退出历<br>史舞台。 | 互联网视频网站点<br>播视频网站CDN   |
| MP4<br>MOV<br>FMP4<br>F4V<br>MPEG-4 Part 14 | MPEG组织<br>基于<br>QuickTim<br>e格式。 | MPEG2<br>MPEG4<br>H. 263<br>H. 264<br>H. 265<br>AV1 | MP3<br>AAC<br>OPUS  | HTTP (TCP) HLS (TCP) DASH (TCP) | 持续更新相关标准,基本跟上<br>高清这一波潮流,发展动力强<br>劲。                        | 互联网视频网站点<br>播视频网站短视频   |
| RAW   | 登虹                               | H. 264<br>H. 265<br>AV1                             | G711<br>AAC<br>OPUS | Tcpbuffer(TCP)                  | 公司私有个数数据,仅仅局限<br>登虹自己产品。                                    | 安防IoT                  |

MPEG是运动图像专家组(Moving Picture Experts Group)的简称,其实质上的名称为国际标准化组织(ISO)和国际电工委员会(IEC)联合技术委员会(JTC)1的第29分委员会的第11工作组,即ISO/IEC JTC1/SC29/WG11,成立于1988年。

### MEPG2基本概念

### 基本概念:

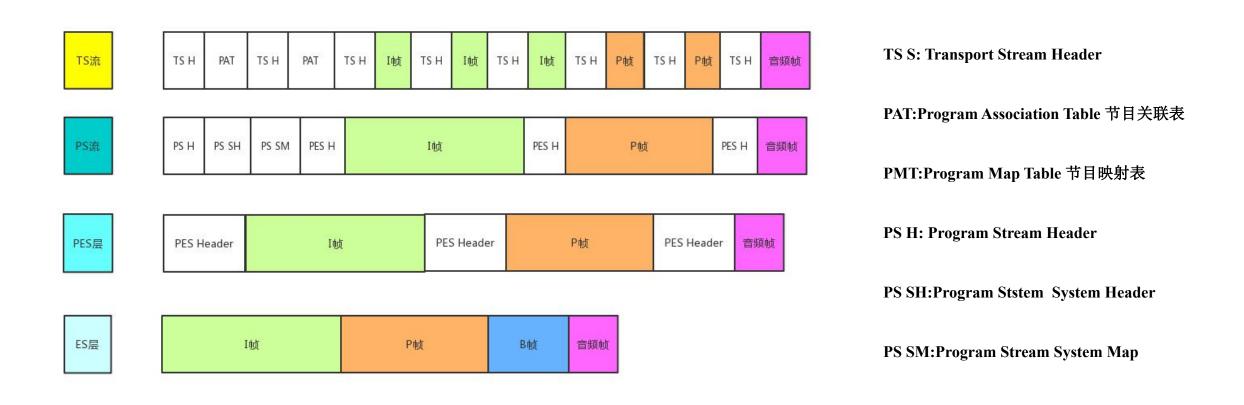
ES即Elemenyary Stream:指只包含1个信源编码器的数据流,就是编码后的原始视频数据或者音频数据;

**PES**即Packetized Elementary Stream:将ES流分组打包后,在每个包前面加PES包头就构成了PES,视频PES一般一帧一个包,音频PES一般一个包的数据量不超过64KB;

**PS流**(Program Stream):节目流, 将具有共同时间基准的一个或多个PES组合(复合)而成的单一数据流 (用于播放或编辑系统,如m2p);

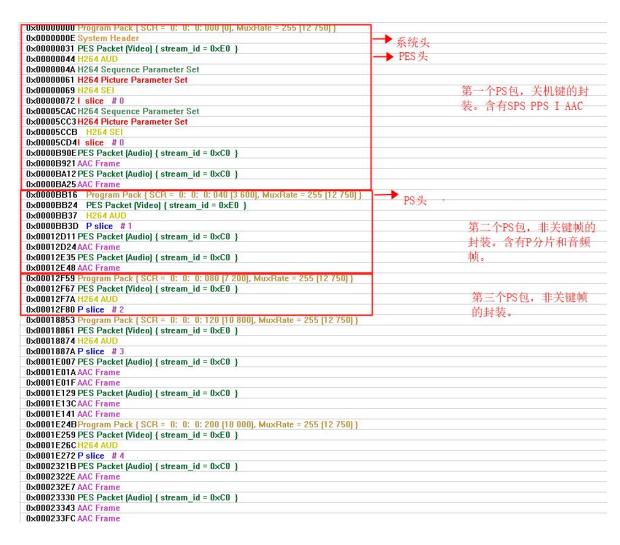
**TS流**(Transport Stream):传输流,将具有共同时间基准或独立时间基准的一个或多个PES组合(复合) 而成的单一数据流(用于数据传输);

## MPEG2码流层次结构



## PS 封装格式解析

## PS码流分析示意图



#### 1. IDR关键帧封装:

每个IDR NALU 前一般都会包含SPS、PPS 等NALU,因此将SPS、PPS、IDR 的NALU 封装为一个PS 包:

PSheader | PS system header | PS system Map | PES header | h264 raw data.

#### 2. 非关键帧封装:

PS包从外到内的顺序是:

PSheader | PES header | h264 raw data.

#### 3. 含有音频的PS封装:

PS头|PES(video)|PES(audio)

PSheader | PES header | h264 raw data | PES header | AAC raw data

### PS Header

marker\_bit

| 语法                                | 位数 | 助记符    | marker_bit   | 1              | bslbf  |       |
|-----------------------------------|----|--------|--|----------------|--|-------|
| pack_header() {                   |    |        | reserved   | 5              | bslbf  | 总     |
| pack_start_code                   | 32 | bslbf  | pack_stuffing_length   | 3              | uimsbf   | 1.    |
| '01'                              | 2  | bslbf  | for (i=0;i <pack_stuffing_length;i++){< td=""><td></td><td></td><td>برم</td></pack_stuffing_length;i++){<> |                |  | برم   |
| system_clock_reference_base[3230] | 3  | bslbf  | stuffing_byte  | 8              | bslbf  | 段     |
| marker_bit                        | 1  | bslbf  | }  |                |  | 2.    |
| system_clock_reference_base[2915] | 15 | bslbf  | if (nextbits()==system_header_start_code)  |                |  |       |
| marker_bit                        | 1  | bslbf  | {  |                |  | p     |
| system_clock_reference_base[140]  | 15 | bslbf  | system_header()  |                |  | 字     |
| marker_bit                        | 1  | bslbf  | }  |                |  |       |
| system_clock_reference_extension  | 9  | uimsbf | }  |                |  | 3.    |
| marker_bit                        | 1  | bslbf  | 0 01 02 03 04 05 06 07 08  |                | The second secon | BCDEF |
| program_mux_rate                  | 22 | uimsbf | 0x0000000  | 7F E0 E8 00 00 | 0 00 01  | <br>c |
| 1980/1914 (4.14.)                 |    |        | 0x0000030 58 00 00 01 E0 B8 D7 80 C0   | OA 31 00 01 00 | 01 11 X1   |       |

总结:

- 1. ps头无论是什么样的包都是要有的,00 00 01 BA是关键,其它字 段基本可以忽略不计;
- 2. 无扩展字段是直接读十四个字节跳过解析即可,但是当 pack\_stuffing\_length不为0时,需要解析出来,加上该长度跳过。该 字段位于第十四字节的后面三bit位;
- 3. 也可以根据该值进行打包和封装,基本对应位都是用默认值;

## PS System Header

| 语 法                          | 位数 | 助记符    |
|------------------------------|----|--------|
| system_header() {            |    |        |
| system_header_start_code     | 32 | bslbf  |
| header_length                | 16 | uimsbf |
| marker_bit                   | 1  | bslbf  |
| rate_bound                   | 22 | uimsbf |
| marker_bit                   | 1  | bslbf  |
| audio_bound                  | 6  | uimsbf |
| fixed_flag                   | 1  | bslbf  |
| CSPS_flag                    | 1  | bslbf  |
| system_audio_lock_flag       | 1  | bslbf  |
| system_video_lock_flag       | 1  | bslbf  |
| marker_bit                   | 1  | bslbf  |
| vedio_bound                  | 5  | uimsbf |
| packet_rate_restriction_flag | 1  | bslbf  |
| reserved_bits                | 7  | bslbf  |
| while (nextbits()=='1') {    |    |        |
| stream_id                    | 8  | uimsbf |
| '11'                         | 2  | bslbf  |
| P-STD_buffer_bound_scale     | 1  | bslbf  |
| P-STD_buffer_size_bound      | 13 | uimsbf |
| }                            |    |        |
| }                            |    |        |

|           | 00 | 0  | 02 | 03 | 04 | 05 | 06 | 07 | 08 | 09 | 0A | 0B | 0C | 0D | 0E | OF | 0123456789ABCDEF                       |
|-----------|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|--|
| 0x0000000 | 00 | 00 | 01 | BA | 44 | 00 | 04 | 00 | 04 | 01 | 00 | 03 | FF | F8 | 00 | 00 | D                                      |
| 0x0000010 | 01 | ВВ | 00 | 09 | 81 | 86 | A1 | 05 | El | 7F | E0 | E8 | 00 | 00 | 00 | 01 | ······································ |
| 0x0000020 | BC | 00 | 0E | E0 | FF | 00 | 00 | 00 | 04 | 1B | E0 | 00 | 00 | 43 | 80 | 9B |  |
| 0x0000030 | 58 | 00 | 00 | 01 | E0 | В8 | D7 | 80 | C0 | 0A | 31 | 00 | 01 | 00 | 01 | 11 | X                                      |
| 0x0000040 | 00 | 01 | 00 | 01 | 00 | 00 | 00 | 01 | 09 | 30 | 00 | 00 | 00 | 01 | 67 | 42 | gB                                     |
| 0x0000050 | 00 | 1F | 95 | A8 | 14 | 01 | 6E | 84 | 00 | 00 | 1F | F4 | 00 | 05 | 7E | 40 | n@                                     |
| 0x0000060 | 10 | 00 | 00 | 00 | 01 | 68 | CE | 3C | 80 | 00 | 00 | 00 | 01 | 06 | E5 | 01 | h.<                                    |
| 0x0000070 | 2F | 80 | 00 | 00 | 00 | 01 | 65 | В8 | 00 | 00 | 6B | 94 | 7C | 46 | 1E | 00 | /ek. F                                 |
| 0800000x0 | 01 | 00 | 09 | 07 | 00 | 1A | 4C | 0B | 28 | 75 | 82 | 30 | 18 | 64 | 0E | B2 | L. (u.0.d                              |
| 0x0000090 | A7 | C1 | 48 | 70 | 86 | 84 | FO | 64 | 03 | 41 | F6 | 5F | FD | BF | F8 | 20 | Hpd.A                                  |
| 0x00000A0 | FC | 2F | 3F | FB | F5 | 79 | DA | 3F | 2F | FF | FA | 78 | 53 | 29 | F7 | FF | ./?y.?/xS)                             |
| 0x00000B0 | D7 | DA | 23 | 56 | 8F | CD | FF | FD | 3E | F1 | 12 | 09 | 33 | 77 | AB | C5 | <b>#</b> ∀>3w                          |
| 0x00000C0 | DE | 2E | 7C | 5D | E7 | C5 | ЗА | BC | 5D | E2 | 99 | F5 | AB | C5 | CF | 89 | ]:.]                                   |
| 0x00000D0 | 3E | Α7 | CD | EA | F1 | 77 | 89 | 0F | 3E | 27 | EA | 7C | 2F | 40 | F7 | A9 | >w>'. /@                               |
| 0x00000E0 | F1 | 27 | D4 | F8 | 5E | 82 | A7 | CD | EA | F1 | 21 | E7 | C4 | 9F | 53 | E1 | .'^!s.                                 |
| 0x00000F0 | 70 | A0 | AB | C1 | A8 | 90 | 1F | 06 | A2 | 40 | F3 | E0 | D4 | 48 | 1E | 7C | p                                      |
| 0x0000100 | 1A | 84 | 82 | A7 | C1 | A8 | 48 | 2F | 3E | 0D | 42 | 41 | 57 | 83 | 50 | 90 | H/>.BAW.P.                             |
| 0x0000110 | 3C | 1A | 89 | 01 | FO | 6A | 12 | 0B | DE | 0D | 44 | 81 | E0 | B2 | 24 | OF | <jd\$.< td=""></jd\$.<>                |
| 0x0000120 | 83 | 50 | 90 | 3D | E0 | AA | 28 | 43 | Cl | A8 | 48 | OF | 83 | 50 | 90 | 74 | .P.=(CHP.t                             |
| 0x0000130 | F8 | 35 | 09 | 07 | 5E | OD | 49 | 01 | FO | 5D | OA | 10 | 3B | A2 | 35 | F4 | .5^.I];.5.                             |
| 0x0000140 | FD | AF | BE | FB | EF | BE | F9 | F0 | 61 | 03 | 40 | 6F | 3E | 0C | 83 | 41 | a.@o>A                                 |

#### 总结:

对于系统头部的解析,我们一般只要先首先判断是否存在系统头(根据系统头的起始码 0x000001BB),然后我们读取系统头的头部长度,即PS SYSTEM HEADER LENGTH部分, 然后根据头部的长度,跳过PS系统头。进入下一个部分,即PS 节目流映射头。

### PS System Map Header

| 语 法  | 位数 | 助记符    |
|--|----|--------|
| program_stream_map() {                                       |    |        |
| packet_start_code_prefix                                     | 24 | bslbf  |
| map_stream_id  | 8  | uimsbf |
| program_stream_map_length                                    | 16 | uimsbf |
| current_next_indicator                                       | 1  | bslbf  |
| reserved   | 2  | bslbf  |
| program_stream_map_version                                   | 5  | uimsbf |
| reserved   | 7  | bslbf  |
| marker_bit   | 1  | bslbf  |
| program_stream_info_length                                   | 16 | uimsbf |
| for (i=0;i <n;i++){< td=""><td></td><td></td></n;i++){<>     |    |        |
| descriptor()   |    |        |
| }  |    |        |
| elementary_stream_map_length                                 | 16 | uimsbf |
| for (i=0;i <n1;i++){< td=""><td></td><td></td></n1;i++){<>   |    |        |
| stream_type  | 8  | uimsbf |
| elementary_stream_id   | 8  | uimsbf |
| elementary_stream_info_length                                | 16 | uimsbf |
| for (i=0;i <n2;i++) td="" {<=""><td></td><td></td></n2;i++)> |    |        |
| descriptor()   |    |        |
| }  |    |        |
| }  |    |        |
| CRC_32   | 32 | rpchof |
| }  |    |        |

Stream Type:

1、MPEG-4 视频流: 0x10; 2、H.264 视频流: 0x1B 3、SVAC 视频流: 0x80 4、G.711 音频流: 0x90 5、G.722.1 音频流: 0x92 6、G.723.1 音频流: 0x93 7、G.729 音频流: 0x99 8、SVAC音频流: 0x9B

4字节描述PS基本流的 音视频编码类型

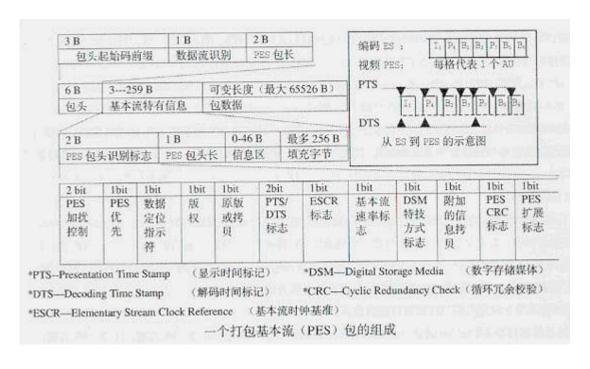
elementary\_sream\_id: 这个字段的定义,其0x(C0~DF) 指音频, 0x(E0~EF)为视频 该基本流所在PES分组的PES 分组标题中stream\_id字段的 值。

|           | 0.0 | 01 | 02 | 03 | 04 | 05 | 06 | 07 | 08 | 09 | 0A | 0B | 0C | 0D | 0E | 0F | 0123456789ABCDEF        |
|-----------|-----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|-------------------------|
| 0x0000000 | 00  | 00 | 01 | BA | 44 | 00 | 04 | 00 | 04 | 01 | 00 | 03 | FF | F8 | 00 | 00 | D                       |
| 0x0000010 | 01  | BB | 00 | 09 | 81 | 86 | A1 | 05 | E1 | 7F | E0 | E8 | 00 | 00 | 00 | 01 |                         |
| 0x0000020 | ВС  | 00 | 0E | E0 | FF | 00 | 00 | 00 | 04 | 18 | E0 | 00 | 00 | 43 | 80 | 9B |                         |
| 0x0000030 | 58  | 00 | 00 | 01 | E0 | В8 | D7 | 80 | C0 | 0A | 31 | 00 | 01 | 00 | 01 | 11 | X1                      |
| 0x0000040 | 00  | 01 | 00 | 01 | 00 | 00 | 00 | 01 | 09 | 30 | 00 | 00 | 00 | 01 | 67 | 42 | gB                      |
| 0x0000050 | 00  | 1F | 95 | A8 | 14 | 01 | 6E | 84 | 00 | 00 | 1F | F4 | 00 | 05 | 7E | 40 | n                       |
| 0x0000060 | 10  | 00 | 00 | 00 | 01 | 68 | CE | 3C | 80 | 00 | 00 | 00 | 01 | 06 | E5 | 01 | h.<                     |
| 0x0000070 | 2F  | 80 | 00 | 00 | 00 | 01 | 65 | В8 | 00 | 00 | 6B | 94 | 7C | 46 | 1E | 00 | /ek. F                  |
| 0x0000080 | 01  | 00 | 09 | 07 | 00 | 1A | 4C | 0B | 28 | 75 | 82 | 30 | 18 | 64 | 0E | B2 | L.(u.0.d                |
| 0x0000090 | A7  | Cl | 48 | 70 | 86 | 84 | F0 | 64 | 03 | 41 | F6 | 5F | FD | BF | F8 | 20 | Hpd.A                   |
| 0x00000A0 | FC  | 2F | 3F | FB | F5 | 79 | DA | 3F | 2F | FF | FA | 78 | 53 | 29 | F7 | FF | ./?y.?/xS)              |
| 0x00000B0 | D7  | DA | 23 | 56 | 8F | CD | FF | FD | 3E | F1 | 12 | 09 | 33 | 77 | AB | C5 | <b>∦</b> V>3w           |
| 0x00000C0 | DE  | 2E | 7C | 5D | E7 | C5 | 3A | BC | 5D | E2 | 99 | F5 | AB | C5 | CF | 89 | ]:.]                    |
| 0x00000D0 | 3E  | A7 | CD | EA | F1 | 77 | 89 | 0F | 3E | 27 | EA | 7C | 2F | 40 | F7 | A9 | >w>'. /@                |
| 0x00000E0 | F1  | 27 | D4 | F8 | 5E | 82 | A7 | CD | EA | Fl | 21 | E7 | C4 | 9F | 53 | E1 | .'^!S.                  |
| 0x00000F0 | 70  | A0 | AB | Cl | A8 | 90 | 1F | 06 | A2 | 40 | F3 | E0 | D4 | 48 | 1E | 7C | p@H.                    |
| 0x0000100 | 1A  | 84 | 82 | A7 | C1 | A8 | 48 | 2F | 3E | 0D | 42 | 41 | 57 | 83 | 50 | 90 | H/>.BAW.P.              |
| 0x0000110 | 3C  | 1A | 89 | 01 | FO | 6A | 12 | 0B | DE | 0D | 44 | 81 | EO | B2 | 24 | OF | <jd\$.< td=""></jd\$.<> |
| 0x0000120 | 83  | 50 | 90 | 3D | E0 | AA | 28 | 43 | Cl | A8 | 48 | OF | 83 | 50 | 90 | 74 | .P.=(CHP.t              |
| 0x0000130 | F8  | 35 | 09 | 07 | 5E | OD | 49 | 01 | FO | 5D | OA | 10 | 3B | A2 | 35 | F4 | .5^.I];.5.              |
| 0x0000140 | FD  | AF | BE | FB | EF | BE | F9 | FO | 61 | 03 | 40 | 6F | 3E | 0C | 83 | 41 | a.@o>A                  |
| 0x0000150 | 79  | F0 | 3C | 80 | 34 | 17 | 9F | 06 | 40 | D0 | 54 | F8 | 07 | D0 | 83 | 9C | y.<.4@.T                |
| 0x0000160 | AB  | C2 | 00 | 83 | AC | 78 | 24 | 43 | 44 | 70 | 12 | 24 | 03 | A8 | D7 | 81 | x*CDI.\$                |

对于这个字段的解析,我们需要取值0x000001BC的位串,指出节目流映射的开始, 暂时不需要处理,读取Header Length直接跳过即可,如果需要解析流编码类型,必 须详细解析这个字段。

实际上我们需要在解析节目映射流头字段时,需要解析基本映射流长度字段,这里面包含这是否含有音视频以及音视频对应的编码格式。

### PES Header



### 重点字段

- 1.Stream id:区分视频还是音频以及私有数据的关键;
- 2.PTS\_DTS\_flags:2位字段。当值为'10'时,PTS字段应出现在PES分组标题中;当值为'11'时,PTS字段和DTS字段都应出现在PES分组标题中;当值为'00'时,PTS字段和DTS字段都不出现在PES分组标题中。值'01'是不允许的;

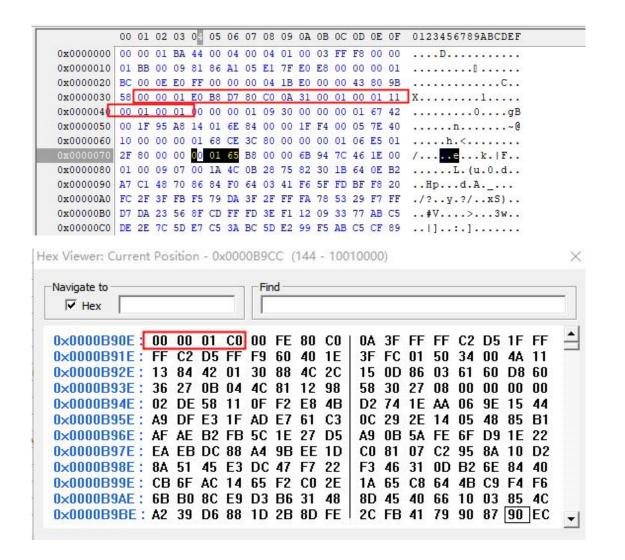
### 总结:

PES显然是里面所有头最复杂的,但是很多字段到底有没有定义,需

要看第八个字节的六大标记,如果没有此标记解析起来相对简单。

- 1. 先解析头的前缀四字节,根据stream id判断是音频还是视频;
- 2.紧接着长度,则为下一个ps包的起始位置;
- 3.判断PTS DTS flags PTS DTS标记;
- 4.解析PTS和DTS值即可;

### PES Header



总结:

PES显然是里面所有头最复杂的,但是很多字段到底有没有定义,需要看第八个CO字节的六大标记,如果没有此标记解析起来相当简单。特别是PES扩展字段标记和特技方式字段。再后面是长度OA,后面的字节则就是PES重要信息PTS和DTS。当然有没有DTS还是要看前面的标记。一种时间戳占五个字节。

- 1. 先解析头的前缀四字节,根据stream\_id判断是音频还是视频;
- 2.紧接着长度,则为下一个ps包的起始位置;
- 3.判断PTS DTS flags PTS DTS标记,解析PTS和DTS值即可;
- 4.stream id里面含有一些私有数据0xBD,实现是需要去掉;

## 国标PS流

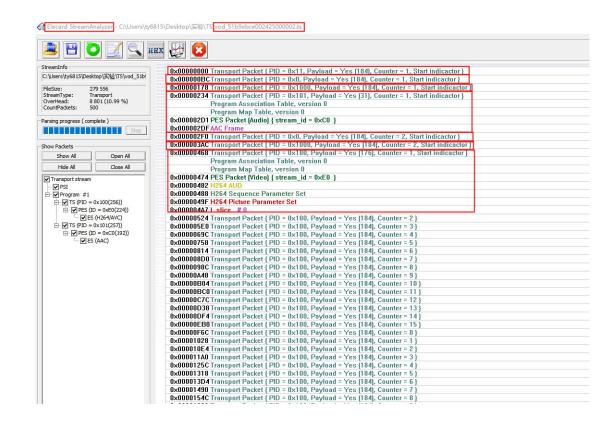
#### 总结:

国标PS流虽然非常复杂,每个包头涉及字段很多。但是实际解析起来可以忽略大量字段,t同时如果是封装,对于不关心字段也就填充默认值来处理:

- 1.PS头即一般就是14字节,直接识别跳过即可;
- 2.系统标题头也是识别前缀,直接读取长度跳过即可;
- 3.节目映射头只有当PES里面承载的ES是IDR帧时才会存在,一般要解析是否还有音视频,同时了解他们的编码格式;
- 4.PES头虽然复杂,但是我们只解析里面的PTS和DTS,里面的六大标记字段只有是0,解析起来也是非常简单的,其中第7和第8字节是关键;

TS 封装格式解析

## TS流分析



PAT Program Association Table:节目关联映射表 PMT Program Map Table:节目映射表

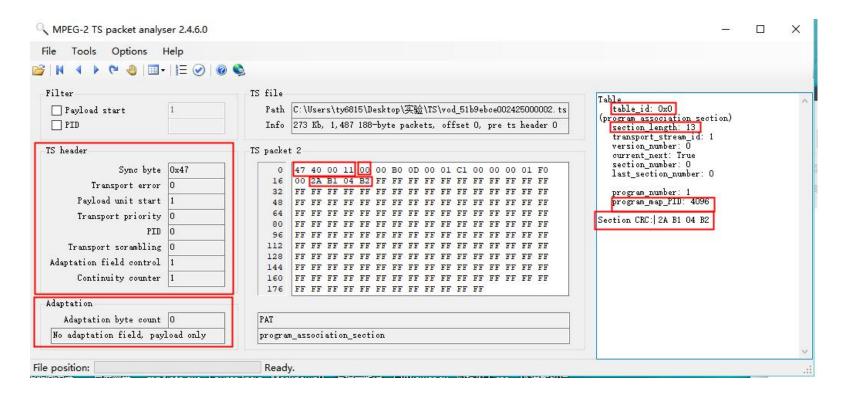
#### 总结:

- 1.TS流由一个个TS包组成,正常情况下没个TS包固定大小是 188字节;
- 2.TS包类型有PAT、PMT、视频包、音频包等,其中PAT是分析TS的起始位置和关键:
- 3.TS流是下面这种形式包:

### PAT PMT PMT DATA DATA .....PAT PMT PMT PMT DATA DATA.....

- 4.TS流和PS流在PES层面是统一的,可以自由相互转换;
- 5.一个完整的PES包由整数个TS包组成,否则需要对最后一个TS进行填充;

### TS Header



#### 重点字段解析:

Payload unit start:表示这个TS包是不是所承载内容的第一个TS包标识承载内容的开始和结束;

Continuity Counter:根据分别承载 内容的不同,分别计数0-15之间 循环;

PID:代表当前TS包承载内容类型 一般第一个TS包的PID是0x00代表 PAT;

## PAT

| 字段                      | 字段含义    | 占位 | 属于第几字节 | 当前数据和分析         |
|-------------------------|---------|----|--------|-----------------|
| table_id                | 表ID     | 8  | 第一字节   | 0x00: 标识一个      |
|                         |         |    |        | TS PSI分段的内      |
|                         |         |    |        | 容是节目关联分         |
|                         |         |    |        | 段,条件访问分         |
|                         |         |    |        | 段还是节目映射         |
|                         |         |    |        | 分段。对于PAT        |
|                         |         |    |        | 置为0x00          |
| section_synatx_indicati | 段同步标识   | 1  | 第二字节   | 1:              |
| on                      |         |    |        |                 |
| "0"                     | 0值      | 1  | 第二字节   | 0:              |
| reserved                | 保留值     | 2  | 第二字节   | 11:             |
| section_length          | 分段长度    | 12 | 第三字节   | 0000 0000 1100: |
|                         |         |    |        | 0x0D: 十三字节。     |
|                         |         |    |        | 分段长度字段,         |
|                         |         |    |        | 这个值是包括该         |
|                         |         |    |        | 字段在内到           |
|                         |         |    |        | CRC_32校验字段      |
|                         |         |    |        | 的字节数,其值         |
|                         |         |    |        | 不超过1021         |
| transport_stream_id     | 传输流标识   | 16 | 第四、五字节 | 0x00 0x01:该字    |
|                         |         |    |        | 节充当标签,标         |
|                         |         |    |        | 识网络内此传输         |
|                         |         |    |        | 流有别于任何其         |
|                         |         |    |        | 他路复用流。其         |
|                         |         |    |        | 值由用户规定。         |
| reserved                | 保留值     | 2  |        | 11:             |
| version_number          | PAT的版本号 | 5  | 第六字节   | 00 000:PAT的版    |
|                         |         |    |        | 本号,如果PAT        |
|                         |         |    |        | 有变,则版本号         |
|                         |         |    |        | 加1              |
|                         |         |    |        | 有变,则版本          |

| current_next_indicator     | 标识        | 1  | 第六字节         | 1:置为0时,表<br>明该传送的段不<br>能使用,下一个<br>表分段才能有效,<br>一般默认值用1     |
|----------------------------|-----------|----|--------------|---|
| section_number             | 分段号       | 8  | 第七字节         | Ox00表明该TS包属于PAT的第几个分段,分段号从0开始。因为PAT可以描述很多PMT信息,所以长度可能比较长。 |
| last_section_number        | 最后一个分段号   | 8  | 第八字节         | 0x00表明该PAT<br>的最大分段数目,<br>一般情况都是一<br>个PAT表由一个<br>TS包传输。   |
| program_number(循环开始,N从O开始) | 节目的编号     | 16 | N+2字节        | 0x00 0x01   |
| reserved                   | 保留值       | 3  | N+3字节        | 111:  |
| network_PID                | NIT表的PID值 | 13 | N+3 N+4字节    | 节目号为0则用 此值;   |
| program_map_PID            | PMT的ID值   | 13 | N+4字节        | 1 0000 0000<br>0000: 0x10<br>0x004096 其它<br>时,则填充此值;      |
| crc_32                     | CRC校验     | 32 | 最后载荷四个字<br>节 | CRC校验   |

### PAT

#### 总结:

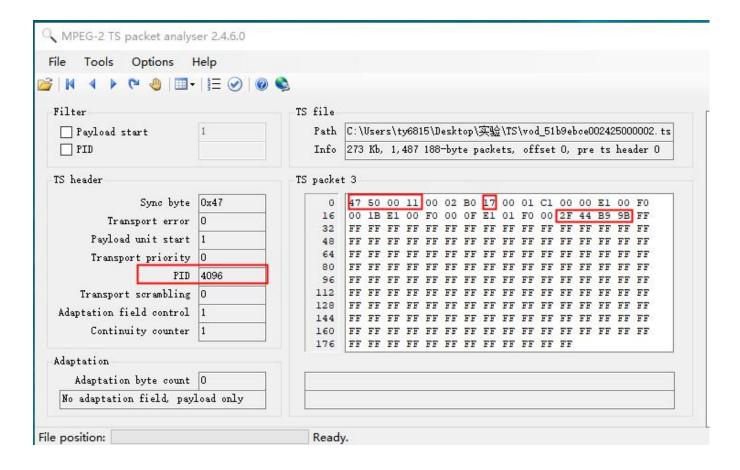
1. 表格灰色部分是个循环,整个占用四字节,那到底有几个循环是怎么算出来的?

是根据section\_length-6-4字节/4算来的,减去6就是section\_length到last\_section\_number字段和,

减去4是因为还有CRC\_32占用的四字节,总长度是0x0d即(13 - 6 -4)/4;

- 2. 表格红色部分,是if-else if关系。取决于前面字段节目号是否从0开始;
- 3. PAT表我们主要就是解析怎么获取到后面的PMT的表格ID;

### **PMT**



- 1. 当前频道中包含的Video数据的PID;
- 2. 当前频道中包含的Audio数据的PID;
- 3. 和当前频道关联在一起的其它数据PID;

## PMT

| 字段                        | 字段含义    | 占位 | 属于第几字节 | 当前数据和分析  |
|---------------------------|---------|----|--------|--|
| table_id                  | 表ID     | 8  | 第一字节   | 0x02: 标识一个TS PSI<br>分段的内容是节目关联<br>分段,条件访问分段还<br>是节目映射分段。对于<br>PMT置为0x02                                |
| section_synatx_indication | 段同步标识   | 1  | 第二字节   | 1:对于PMT该字段置为1  |
| "0"                       | 0值      | 1  | 第二字节   | 0:   |
| reserved                  | 保留值     | 2  | 第二字节   | 11:  |
| section_length            | 分段长度    | 12 | 第二、三字节 | 0000 0001 0111: 0x17: 23: 二十三字节。分段<br>长度字段,前两位置00,<br>这个值是包括该字段在<br>内到CRC_32校验字段的<br>字节数,起值不超过<br>1021 |
| program_number            | 传输流标识   | 16 | 第四、五字节 | 0x00 0x01:对应于PAT中<br>的program_number   |
| reserved                  | 保留值     | 2  | 第六字节   | 11:  |
| version_number            | PMT的版本号 | 5  | 第六字节   | 00 000:PMT的版本号,<br>如果字段中有关信息有<br>变,则版本号以32为模<br>加1。版本号是对一个节<br>目的定义。                                   |
| current_next_indicator    | 标识      | 1  | 第六字节   | 1:该字段置为1时,表示<br>当前传送的<br>program_map_section可<br>用。置为0时,表明该传<br>送的段不能使用,下一<br>个表分段才能有效;                |

| section_number                      | 分段号     | 8  | 第七字节           | 0x00该字段一般总是<br>置为0x00   |
|-------------------------------------|---------|----|----------------|---|
| last_section_number                 | 最后一个分段号 | 8  | 第八字节           | 0x00该字段一般总是<br>置为0x00   |
| reserved                            | 保留号     | 3  | 第九字节           | 1110 0001 0000 0000   |
| PCR_PID                             | PCR值    | 13 | 第九字节第十<br>字节   | 0 0001 0000<br>00000x100256该字段<br>指示TS包的PID值,该<br>TS含有该PCR字段,而<br>PCR值对应于有节目号<br>指定的节目。           |
| reserved                            | 保留值     | 4  | 第十一字节          | 1111  |
| program_info_length不<br>为0,后面进行循环层1 | 节目信息长度  | 12 | 第十一字节第<br>十二字节 | 0000 0000 000: 表明<br>跟随其后的对节目信息描述的字节数,也<br>就是第一个N loop<br>descriptors的字节数。<br>这里是0则表示第一层<br>循环略过; |

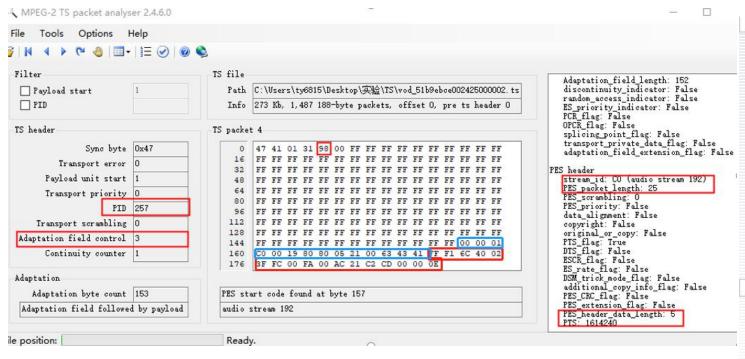
### PMT

| stream_type开始循环层2      | 流类型                  | 8  | N+1字节(N=0)          | 0x1B:表明这个流是h264编码<br>格式;表示PES流的类型。                                    |
|------------------------|----------------------|----|---------------------|---|
| reserved               | 保留值                  | 3  | N+2字节 (N=0)         | 111:  |
| elementary_pid         | 负载该PES流的TS包的<br>PID值 | 13 | N+2字节N+3字节<br>(N=0) | <b>0 0001 0000 0000:</b> 0x100: 256: 表明负载该PES流的PID 值                  |
| reserved               | 保留值                  | 4  | N+4字节 (N=0)         | 1111  |
| es_info_length不为0,循环层3 | Es流描述相关的字节数          | 12 | N+5字节(N=0)          | 0000 0000:<br>表明跟随其后描述相关节目元<br>素的字节数: 否则为第二个循<br>环的第二层循环;             |
| stream_type            | 流类型                  | 8  | N+1字节 (N=1)         | 0x0F  |
| reserved               | 保留值                  | 3  | N+2字节 (N=1)         | 111   |
| elementary_pid         | 负载该PES流的TS包的<br>PID值 | 13 | N+2字节N+3字节<br>(N=1) | 0 0001 0000 0001:0x101257:  |
| reserved               | 保留值                  | 4  | N+4字节(N=1)          | 1111:   |
| es_info_length         | Es流描述相关的字节数          | 12 | N+5字节(N=1)          | 0000 000000000: 表明跟随其<br>后描述相关节目元素的字节数;<br>否则为第二个循环的第二层循<br>环; 默认一般为0 |
| crc_32                 | CRC校验                | 32 | 载荷最后四字节             | 0x2F 0x44 0xB9 0x9B CRC校<br>验   |

#### 总结:

- 1. program\_info\_length如果不为0,则有多少字
- 节,则后面要跟多少字节对节目信息进行描述。
- 2. stream\_type到es\_info\_length是另外一层循环,在这里面有可能还存在一层循环,就是es\_info\_length,不为0时要将此字节计算在之内,如果为0则一次循环需要五字节。
- 3. 图表中显示了红色和黄色循环了两次。
- 4. 这里面定义的真实码流视频和音频的PID,所以PMT是定义每路节目的音视频类型TYPE和编号PID的关键。

## 音频TS



### $0 \times 000002D1$ PES Packet (Audio) { stream id = $0 \times C0$ } packet length = 25 PES\_scrambling\_control = 0 PES priority = 0 data alignment indicator = 0 copyright = 0 original or copy = 0 PTS DTS flags = 2 ESCR flag = 0 ES rate flag = 0 DSM trick mode flag = 0 additional copy info flag = 0 PES CRC flag = 0 PES extension flag = 0 PES header data length = 5 PTS = 0: 0: 17: 936 [1 614 240] 0x000002DFAAC Frame id = 0laver = 0 protection absent = 1 profile = 1 sf index = 11

private bit = 0

original = 0 home = 0

channel configuration = 1

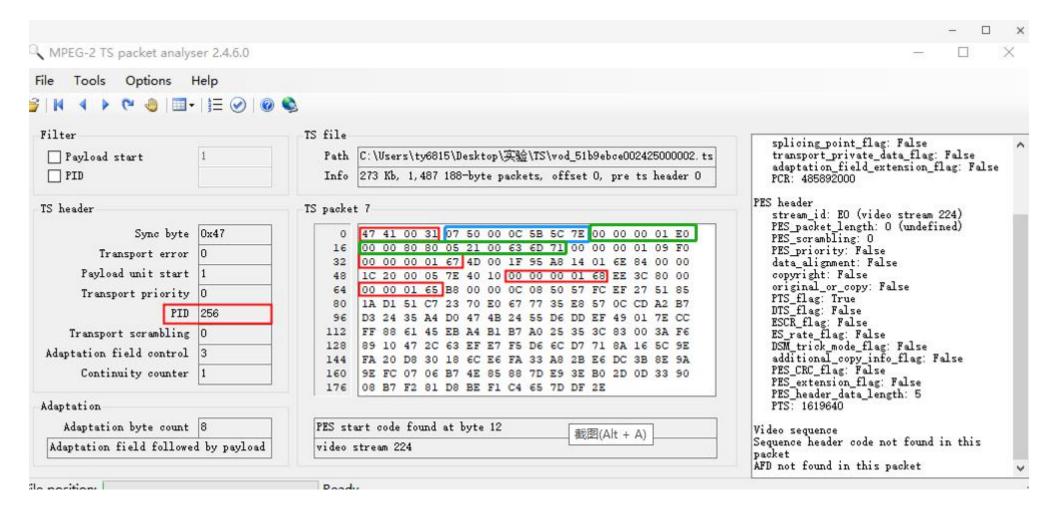
aac\_frame\_length = 17 adts\_buffer\_fullness = 2047 no raw data blocks in frame = 0

SamplingRate = 8000

Channels = 1 Duration = 206258

copyright\_identification\_bit = 0 copyright\_identification\_start = 0

## 视频TS

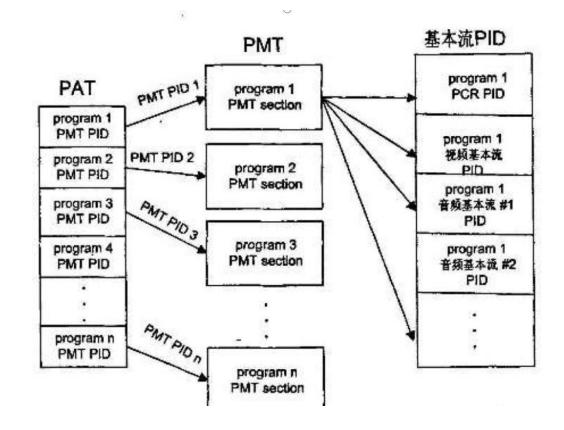


## TS流总结

#### 总结:

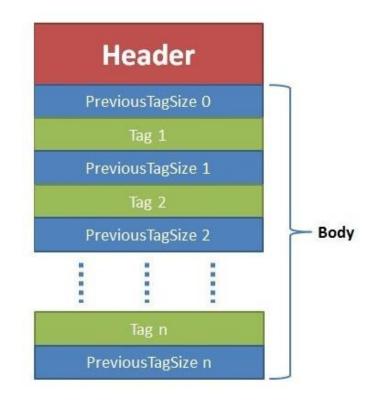
基本上从TS流分析H264码流这块,简单的方法就是先分析PID为0x00的PAT,然后分析PMT里面对音频和视频的PID定义。然后以定义PID分别过滤音频,然后再过滤视频,这样就完成了从TS文件到音视频数据的过滤。

- 1.第一步找PAT.注重分析PMT的表的PID;
- 2.第二步找PMT,分析里面是否含有音视频,音视频编码格式,音视频的PID;
- 3.第三步根据音视频的PID过滤音频TS包和视频TS包
- 4.拿掉TS头,PES头,后即可得到音频数据和视频数据的裸码流;



## FLV 封装格式解析

## FLV整体结构



|            | Signatur  | re (3 字节) 为  | o文件标识,总为"FLV",(0x46, 0x4c, 0x66)  |  |  |  |  |  |  |
|------------|---|--|---|--|--|--|--|--|--|
|            | Version   | (1字节) 为履   | 反本,目前为 0x01   |  |  |  |  |  |  |
| Flv Header | Flags (1字节)前5位保留,必须为0。第6位表示是否存在音频 Tag。第7位保留,必须为0。第8位表示是否存在视频 Tag。 |  |   |  |  |  |  |  |  |
|            |   | Headersize(4字节)为从 File Header 开始到 File Body 开始的字节数,版本 1 中总为 9。 |   |  |  |  |  |  |  |
|            | Previous  | Tag Size #0 (4   | 4 字节)表示前一个 Tag 的长度  |  |  |  |  |  |  |
|            |   |  | Type (1 字节) 表示 Tag 类型,包括音频 (0x08),视频 (0x09) 和 script data (0x12),其他类型值被保留 |  |  |  |  |  |  |
|            |   |  | Datasize(3 字节)表示该 Tag Ddata 部分的大小   |  |  |  |  |  |  |
|            |   | Tag Header   | Timestamp(3 字节)表示该 Tag 的时间戳   |  |  |  |  |  |  |
|            | Tag #1  |  | Timestamp_ex (1 字节)表示时间戳的扩展字节,当 2.<br>位数值不够时,该字节最为最高位将时间戳扩展为 3.<br>位数值    |  |  |  |  |  |  |
|            |   |  | StreamID (3字节)表示 stream id 总是 0   |  |  |  |  |  |  |
| Flv Body   |   | Tag Data   | 不同类型 Tag 的 data 部分结构各不相同,当 header 的结构是相同的                                 |  |  |  |  |  |  |
|            | Previous  | Tag size #1 以  | Tag #1 的大小(11 + Datasize)   |  |  |  |  |  |  |
|            | Tag #2  |  |   |  |  |  |  |  |  |
|            | Previous  | Previous Tag size #2   |   |  |  |  |  |  |  |
|            |   |  |   |  |  |  |  |  |  |
|            | Tag #N  |  |   |  |  |  |  |  |  |
|            | Previous  | Tag size #N  |   |  |  |  |  |  |  |

flv header + flv body

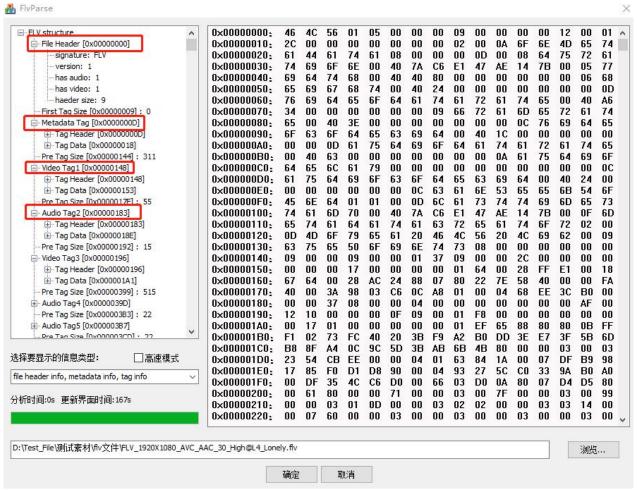
flv header + previous size0 + tag1 + previous size1 + tag2 + .....+ prvious sizen+ tagn+1

flv header + previous size0 + tag1 header + tag1 data1 + ......+ previous sizen + tagn header+tagn data

previous tag sizen 表示的前一个tag的整体大小

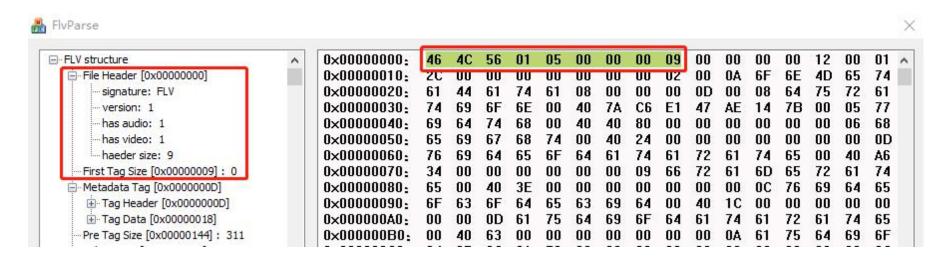
每个Tag数据大小在Tag Header头里面的Data size字段里面;

## 实例分析



| × |  |  |  |  |
|---|--|--|--|--|
|   |  |  |  |  |
| • |  |  |  |  |
|   |  |  |  |  |
|   |  |  |  |  |
|   |  |  |  |  |

### FLV Header



#### 总结:

- 1. Flv header 的前三个字节是固定的FLV的 ASCII 码的值0x46 0x4C 0x56;
- 2. 接下来的一个字节表示 FLV 的版本号.例如 0x01 代表 FLV 版本号为 1:
- 3. 第 5 个字节中的第0位和第2位分别表示video和audio的存在情况(1表示存在,0 表示不存在)其余6位必须为0.最后的4

字节表示FLV Header的长度,对于version 1,此处为9:

- 4. 一般判断格式是不是flv, 先从收到数据的第一字节连续查找flv三个字符, 如果找到一般可以判断是flv封装格式;
- 5. Header头数据一般是9字节但是不绝对,所以需要读最后的长度字段来解析;

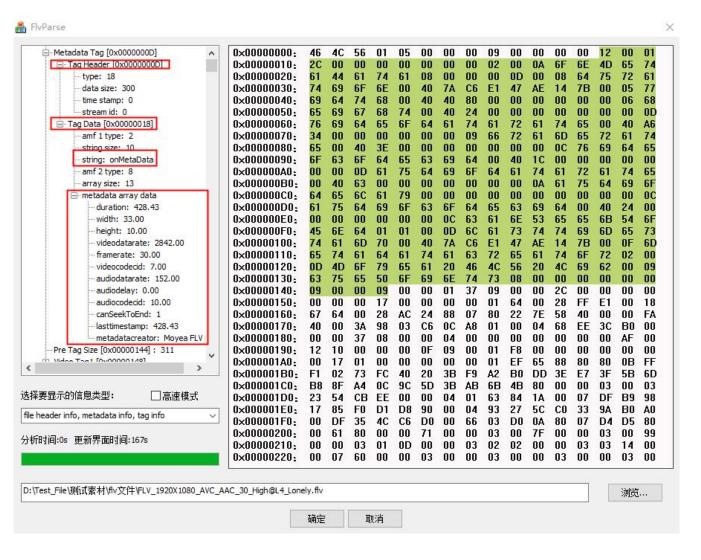
### FLV TAG Header

| 字段                       | 占位           | 备注  |
|--------------------------|--------------|---|
| Tag类型(TagType)           | 1 bytes      | 1-2bit位: 必须为0, 保留位; 第<br>3bit位: 0表示未加密, 1表示加<br>密, 一般默认0; 4-8bit位: 8:<br>音频、9: 视频、18: script数据; |
| 数据大小(DataSize)           | 3 bytes      | 数据字段的长度,是Tag Data的<br>长度,不包括11字节的Tag<br>Header长度;   |
| 时间戳(Timestamp)           | 3 bytes      | 毫秒为单位,第一个tag时,该值总是0,单位是 <mark>毫秒</mark> ,则意味着要将时间戳单位关系换算好;                                      |
| 时间戳扩展(TimeStampExtended) | 1 bytes      | 时间戳扩展为4bytes,代表高8位,<br>很少用到;  |
| 流ID(Stream ID)           | 3bytes       | 总是0,暂时未用到,因为flv只<br>封装一路音视频,但是对于TS和<br>PS则有特殊含义;  |
| 数据(Data)                 | 音频、视频或script | 数据实体,音频、视频和脚本数<br>据;  |

#### 总结:

- 1.无论那种类型的tag,tag头字节都是11字节,要解析里面的音频帧,视频帧或者元数据需要读取tag头里面的data长度字段;
- 2.时间戳很关键,播放过程中,FLV tag的时间信息完全依赖于 FLV 时间戳,内置的其他时间信息都被忽略掉,一般非音视频的tag,时间戳就是0即可;
- 3.注意计算好时间戳大小,这里的单位是毫秒,所以一定要根据采样率和视频帧率,音频帧采样多少计算好时间戳,然后还要换算成毫秒;
- 4.Tag头解析完后,Tag数据部分不是视频帧和音频帧,还要根据H264和AAC的打包方案解析才能得到真实的音视频裸数据;

## Script TAG 脚本元数据TAG



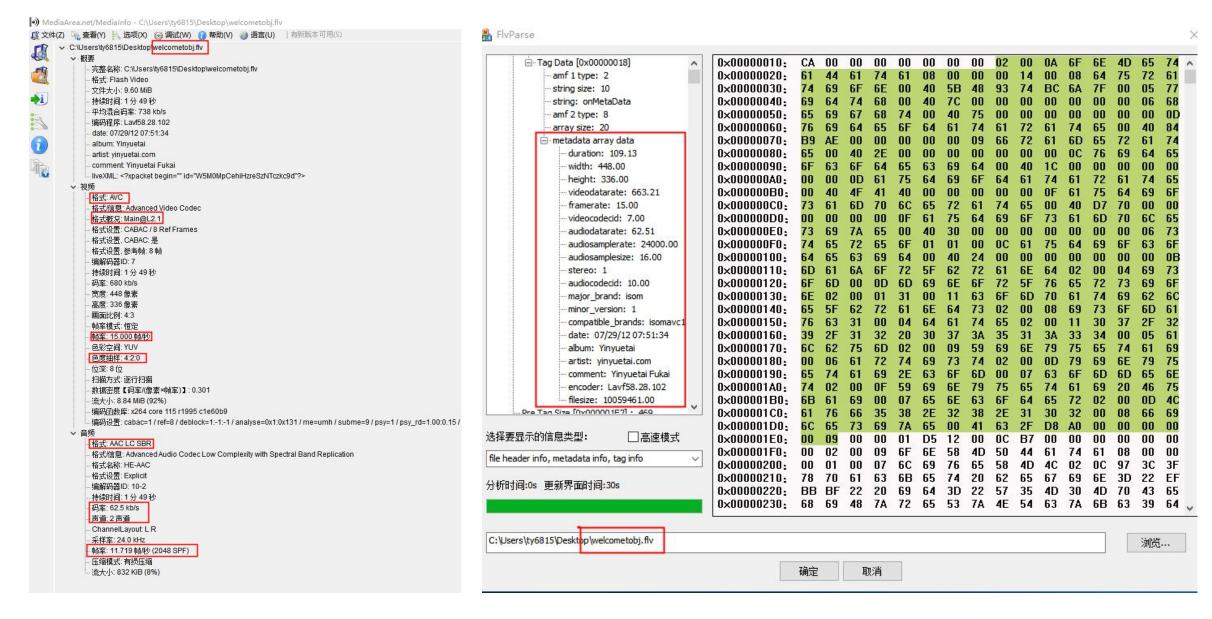
1. Metdata Tag有下面两种包组成:

AMF1{"on MetaData"}|AMF2{"width height"}

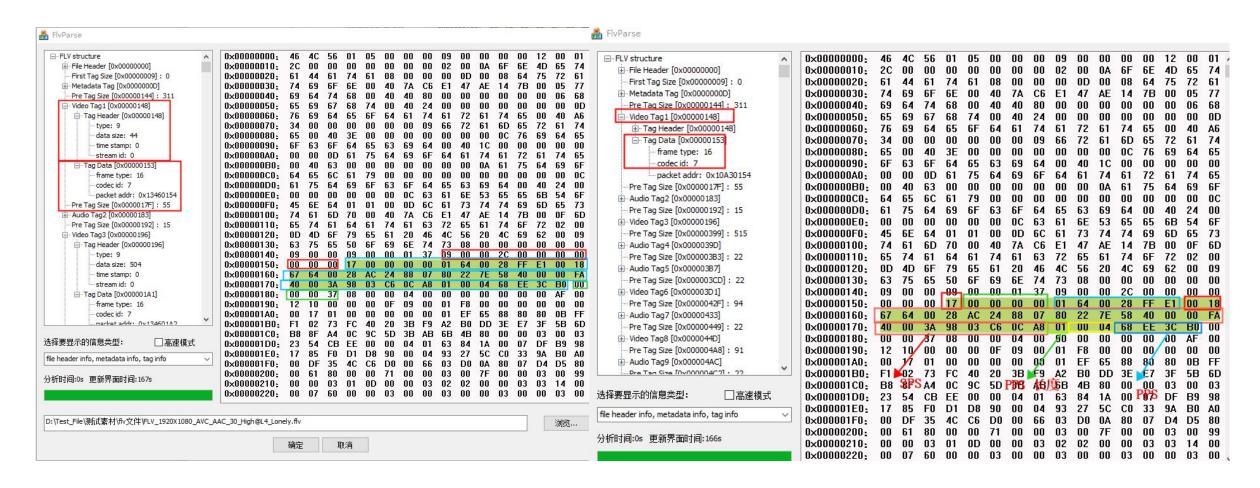
| 字段         | 占位          | 备注  |
|------------|-------------|---|
| AMF Type   | 1byte<br>字节 | 固定值为0x08  |
| AMF 长度     | 4byte<br>字节 | 数组元素的个数:0x00 00 00 0D   |
| AMF value值 | 计算得<br>到字节  | 因为在该tag的头已经说明该tag的data<br>有300字节,则用该值减去AMF1长度13<br>字节和AMF2头的5字节即为剩余AMF2<br>的Value值大小:300 - 13 - 5 = 283字<br>节 |

| 值               | 含义     |
|-----------------|--------|
| duration        | 时长     |
| width           | 视频宽度   |
| height          | 视频高度   |
| videodatarate   | 视频码率   |
| framerate       | 视频帧率   |
| videocodecid    | 视频编码方式 |
| audiosamplerate | 音频采样率  |
| audiosamplesize | 音频采样精度 |
| stereo          | 是否为立体声 |
| audiocodecid    | 音频编码方式 |
| filesize        | 文件大小   |

## Script TAG 脚本元数据TAG

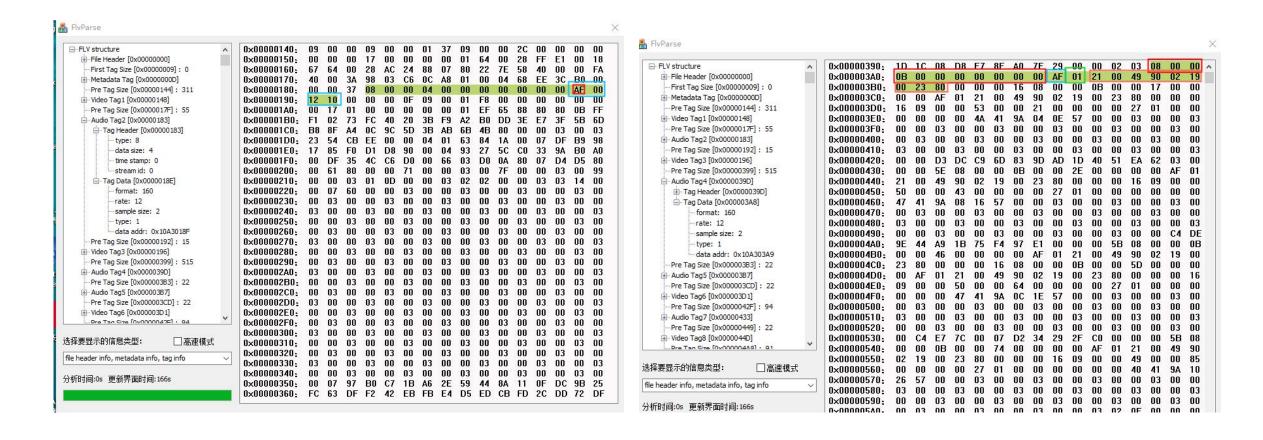


## Video TAG 视频TAG



Video TAG Header Video TAG Data

## Audio TAG 音频TAG



Audio TAG Header

Audio TAG Data

AAC Sequence Header|AAC Raw Data

## FLV总结

- 1.FLV就三种TAG,比较简单的封装格式,每个TAG都是由Tag Header和TAG Data组成;
- 2.FLV中对时间戳的处理?如果含有B帧时间戳怎么计算?特别是PTS和DTS时间戳;
- 3.你能从FLV这种封装格式中,借鉴到什么?加入让你设计一种私有封装格式,又该如何设计;
- 4.如果FLV里面扩展H.265视频编码格式,你觉得大概的思路是什么;
- 5.你觉得FLV的缺点有哪些;

## 参考文档

- 1. ISO/IEC 13818
- 2. ISO/IEC 14496
- 3. video\_file\_format\_spec\_v10\_1.pdf