Video Services Forum (VSF)
Technical Recommendation TR-04

Utilization of ST-2022-6 Media Flows within a VSF TR-03 Environment

November 12, 2015

VSF_TR-04_2015-11-12

INTELLECTUAL PROPERTY RIGHTS

THE POSSIBILITY HAS BEEN RAISED TO THE VSF THAT THE PRACTICE OR IMPLEMENTATION OF THIS RECOMMENDATION MAY INVOLVE THE USE OF A THIRD PARTY INTELLECTUAL PROPERTY RIGHT. HOWEVER, THE OWNER OF SUCH RIGHT HAS NOT AS YET ASSERTED SUCH RIGHT, NOR HAS THE VSF AS YET CONTACTED SUCH THIRD PARTY.

VSF TAKES NO POSITION CONCERNING THE EVIDENCE, VALIDITY OR APPLICABILITY OF THIS OR ANY OTHER INTELLECTUAL PROPERTY RIGHTS, WHETHER ASSERTED BY VSF MEMBERS OR OTHERS OUTSIDE OF THE RECOMMENDATION DEVELOPMENT PROCESS. IF A PATENT HOLDER ASSERTS SUCH RIGHTS AND FILES A STATEMENT WITH THE VSF OF WILLINGNESS TO GRANT A LICENSE UNDER SUCH RIGHTS ON REASONABLE AND NON-DISCRIMINATORY TERMS AND CONDITIONS TO APPLICANTS DESIRING TO OBTAIN SUCH A LICENSE, THEN DETAILS MAY BE OBTAINED FROM THE VSF.

LIMITATION OF LIABILITY

VSF SHALL NOT BE LIABLE FOR ANY AND ALL DAMAGES, DIRECT OR INDIRECT, ARISING FROM OR RELATING TO ANY USE OF THE CONTENTS CONTAINED HEREIN, INCLUDING WITHOUT LIMITATION ANY AND ALL INDIRECT, SPECIAL, INCIDENTAL OR CONSEQUENTIAL DAMAGES (INCLUDING DAMAGES FOR LOSS OF BUSINESS, LOSS OF PROFITS, LITIGATION, OR THE LIKE), WHETHER BASED UPON BREACH OF CONTRACT, BREACH OF WARRANTY, TORT (INCLUDING NEGLIGENCE), PRODUCT LIABILITY OR OTHERWISE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. THE FOREGOING NEGATION OF DAMAGES IS A FUNDAMENTAL ELEMENT OF THE USE OF THE CONTENTS HEREOF, AND THESE CONTENTS WOULD NOT BE PUBLISHED BY VSF WITHOUT SUCH LIMITATIONS.

Executive Summary

Development of IP-based in-studio networking systems and devices exists today in a mixed environment of SDI-based and IP-based equipment. The use of ST 2022-6 for IP-based in-studio networking has attained a degree of use, partially because of the technological simplicity of interworking the signals with SDI. Augmenting the ST 2022-6 signals with AES-67 audio signals and networked metadata allows interoperation with professional audio equipment based on these standards, while still preserving the simplicity of the "multiplexed" nature of SDI where it simplifies the design.

This Technical Recommendation builds upon the system defined in VSF TR-03, and defines SMPTE ST 2022-6 as a video payload within those structures. The resulting method allows for the SDI multiplex, as well as the individual audio and metadata streams, to be individually forwarded through a network.

Table of Contents

1. Introduction

SMPTE 2022-6 documents the transmission of SDI Video Signals (typically with embedded audio and ancillary data) over IP networks. In a typical television production environment, there is a high degree of utility in being able to extract the audio signals from an SDI video signal, process them (mixing, etc) and then to integrate the resulting audio signals back onto a video signal. In parallel to the development of this document, the Video Services Forum Studio Video over IP (SVIP) working group has developed a technical recommendation (VSF TR-03: Transport of Uncompressed Elementary Stream Media over IP) regarding the transport of video, audio, and other signals over IP as separate streams.

This document defines SMPTE ST 2022-6 as a payload type in the context of a VSF TR-03 system. In this use-case, the SMPTE ST 2022-6 stream serves as an alternative type of media flow in addition to those described in TR-03. For example, a device could transmit a ST 2022-6 stream, several AES67 audio streams, and an ANC data stream, described by a common SDP. Additionally, this document specifies a set of "studio-profile" constraints on the ST2022-6/7 when used in this context.

1.1. Contributors

The Chair of the HBRAV activity group which developed this Technical Recommendation was John Mailhot, Imagine Communications. The following individuals participated in the group as well.

Rakesh Jalali (Evertz)	Ali Jerbi (Cisco)	Brian Keane (Aperi)
Bob Edge (Consultant)	Bob Ruhl (VSF)	Brad Gilmer (Gilmer & Associates)
Brent Leroux (Harmonic)	Charles Meyer (GrassValley)	Charlie Yang (Harmonic)
Chin Koh (Nevion)	David Berry (Sony)	Eric Frankhauser (Evertz)
Francois Mace (Intopix)	G. Faubert (Dimetis)	Helge Stephansen (Nevion)
Inge Hillestad (Nevion)	Jack Douglas (Packetstorm)	John Clark (Artel)
John Dale (Medialinks)	Jed Deame (Coherent Video)	Jean-Marie Cloquet (Barco)
Jim Welch (IneoQuest)	Joe Cerra (Macnica)	John Mailhot (Imagine Communications)
Al Kovalick (Consultant)	Leigh Whitcomb (Imagine Communications)	Marc Levy (Macnica)
Mark Alestra	Matt Klein (Xilinx)	Mike Bany (DVBLink)
Kazuki Narita (Medialinks)	Paul Gardiner (Sony)	Philippe Lemonnier (B-Com)
Pierre Costa (AT&T)	Richard Dellacanonica (Artel)	Richard Bullock (Ericsson)
Robert Patton (Medialinks)	Valerie Popie (EVS)	Wes Simpson (Consultant)
Stephen Olday (Sony)	Jun-Ya Suzuki (Panasonic)	Thomas Kernen (Cisco)
Thomas Edwards (Fox)	Tom Sun (Xilinx)	Toshiaki Kojima (Sony)

1.2. About the Video Services Forum

The Video Services Forum, Inc. (www.videoservicesforum.org) is an international association dedicated to video transport technologies, interoperability, quality metrics and education. The VSF is composed of service providers, users and manufacturers. The organization's activities include:

- providing forums to identify issues involving the development, engineering, installation, testing and maintenance of audio and video services;
- exchanging non-proprietary information to promote the development of video transport service technology and to foster resolution of issues common to the video services industry;
- identification of video services applications and educational services utilizing video transport services;
- promoting interoperability and encouraging technical standards for national and international standards bodies.

The VSF is an association incorporated under the Not For Profit Corporation Law of the State of New York. Membership is open to businesses, public sector organizations and individuals worldwide. For more information on the Video Services Forum, contact Bob Ruhl, Operations Manager, Video Services Forum, +1 609 410 6767,

bob.ruhl1@verizon.net.

2. Conformance Notation

Normative text is text that describes elements of the design that are indispensable or contains the conformance language keywords: "shall", "should", or "may". Informative text is text that is potentially helpful to the user, but not indispensable, and can be removed, changed, or added editorially without affecting interoperability. Informative text does not contain any conformance keywords.

All text in this document is, by default, normative, except: the Introduction, any section explicitly labeled as "Informative" or individual paragraphs that start with "Note:"

The keywords "shall" and "shall not" indicate requirements strictly to be followed in order to conform to the document and from which no deviation is permitted.

The keywords, "should" and "should not" indicate that, among several possibilities, one is recommended as particularly suitable, without mentioning or excluding others; or that a certain course of action is preferred but not necessarily required; or that (in the negative form) a certain possibility or course of action is deprecated but not prohibited.

The keywords "may" and "need not" indicate courses of action permissible within the limits of the document.

The keyword “reserved” indicates a provision that is not defined at this time, shall not be used, and may be defined in the future. The keyword “forbidden” indicates “reserved” and in addition indicates that the provision will never be defined in the future.

A conformant implementation according to this document is one that includes all mandatory provisions ("shall") and, if implemented, all recommended provisions ("should") as described. A conformant implementation need not implement optional provisions ("may") and need not implement them as described.

Unless otherwise specified, the order of precedence of the types of normative information in this document shall be as follows: Normative prose shall be the authoritative definition; Tables shall be next; followed by formal languages; then figures; and then any other language forms.

3. References (normative)

SMPTE ST 2022-5:2013 Forward Error Correction for Transport of High Bit Rate Media Signals over IP Networks

SMPTE ST 2022-6:2012 Transport of High Bit Rate Media Signals over IP Networks (HBRMT)

SMPTE ST 2022-7:2013 Seamless Protection Switching of SMPTE ST 2022 IP Datagrams

VSF TR-03 (2015) Transport of Uncompressed Elementary Stream Media over IP

AES67-2013 AES standard for audio applications of networks - High-performance streaming audio-over-IP interoperability

SMPTE ST 2059-1:2015 Generation and Alignment of Interface Signals to the SMPTE Epoch

SMPTE ST 2059-2:2015 SMPTE Profile for Use of IEEE-1588 Precision Time Protocol in Professional Broadcast Applications

IEEE-1588-2008 - IEEE Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems, July 2008, Institute of Electrical and Electronics Engineers (IEEE), US.

RFC 4566 Session Description Protocol, Internet Engineering Task Force

RFC 7273 RTP Clock Source Signaling, Internet Engineering Task Force, June 2014

RFC 7104 Duplication Grouping Semantics in the Session Description Protocol", Internet Engineering Task Force, January 2014

RFC 5888 The Session Description Protocol (SDP) Grouping Framework, Internet Engineering Task Force, June 2010

RFC 6364 Session Description Protocol Elements for the Forward Error Correction (FEC) Framework, Internet Engineering Task Force, October 2011

4. Definitions

Network Epoch is a specific moment in the past at which the Network Clock value was zero. Note that in IEEE-1588-2008, SMPTE ST 2059-1, VSF TR-03, and AES67, the Network Epoch is defined as 1970-01-01T00:00:00TAI.

Network Clock (as defined in AES67) is expressed in seconds (and fractions thereof) referenced to a prevailing network timebase. Note: All participating endpoints strive to have an identical value of the network clock at the same instant in time. The Network Clock is represented with sufficient resolution that it will not overflow during the working life of this document.

2022-6 Media Clock is defined herein as a 27 MHz clock, frequency locked to the ***Network Clock***, and referenced to the ***Network Epoch***.

2022-6 RTP Clock is defined in 2022-6 as samples of a 27 MHz clock. In this document it is required that the *2022-6 RTP Clock* be frequency locked to the ***Network Clock***. The lowest 32 bits of the *2022-6 RTP Clock* are transmitted in the 2022-6 RTP packet header. The method of communicating the fixed offset between the *2022-6 RTP Clock* and the ***Network Epoch*** is documented herein.

First 2022-6 Packet is the SMPTE 2022-6 packet immediately following the packet in which the "M" bit (Marker bit) is set. The M bit is defined in 2022-6 to indicate the last packet of a frame, so the packet immediately after is the first packet of the following frame.

5. Interoperation ST-2022-6 Media Flows with VSF TR-03

All ST 2022-6 originating and receiving devices compliant to this document shall support stream transmission and reception via IPV4 multicast addresses. Unicast addressing shall be supported also.

Devices within TR-03 may create and expose an SDP object compliant to RFC4566. This SDP object specifies (among other things) the prevailing network clock source, and also the offset between the elementary stream *RTP Clock* and the *Media Clock*.

For originating and receiving devices compliant to this document, the ***2022-6 RTP Clock*** shall be frequency locked to the Network Clock, and related to the ***Network Epoch*** by a fixed offset. The fixed offset shall be signaled in the manner described in TR-03. All of the pixels of a video frame (or of an interlaced field) shall be considered to have the same Media Clock value for synchronization purposes, even though they are delivered across many packets with increasing RTP timestamp values. The Media Clock of a progressive-scan SDI frame shall be calculated based on the RTP timestamp of the ***First 2022-6 Packet*** of that frame. The Media Clock of the first field of an interlaced SDI frame shall be calculated in the same manner as a progressive frame. The Media Clock of the second field of an interlaced frame shall be calculated as the mid-point between the Media Clock values of the surrounding first fields, rounding down when necessary.

Devices which originate ST 2022-6 video flows and TR-03 audio and metadata flows compliant to this document may create and expose an SDP object detailing (through media sections) the ST 2022-6 stream and any other streams. The session header of the SDP object shall include a ts-refclk attribute defining the source of the ***Network Clock***, as detailed in TR-03:

```
a=ts-refclk:ptp=IEEE1588-2008:39-A7-94-FF-FE-07-CB-D0:0
```

Each ST-2022-6 media section within the SDP shall contain the following required "m=", "c=", and "a=" clauses:

```
m=video <port> RTP/AVP 98
c=IN IP4 <ipv4-addr>/32
a=rtpmap:98 SMPTE2022-6/27000000
```

where <port> is the UDP destination port, and <ipv4-addr> is the destination address of the stream.

If the fixed offset between the ***2022-6 RTP Clock*** and the ***Network Epoch*** is not zero, then originating devices under this document shall include a media clock offset attribute (RFC7273) in the related media section of the SDP:

```
a= mediaclk:direct=<offset>
```

where <offset> is the ascii decimal representation of the 32 least-significant bits of the difference between the ***Epoch*** of the ***2022-6 RTP clock*** and the ***Network Epoch***. Note that the offset in this attribute for 2022-6 streams is defined in 27MHz clock units, while the offset attributes for other streams are defined in the units of the media clocks for those streams.

6. Integration of 2022-5 FEC and 2022-7 Protection into SDP

If ST-2022-7 seamless merge protection is used across a pair of 2022-6 streams, then the SDP session header shall include an "a=group:DUP <xxx> <yyy>" attribute as defined in RFC 7104

section 4.2, and each 2022-6 related media section shall contain an "a=mid:<xxx>" attribute linking to the grouping attribute (see also RFC 5888). For example:

```
a=group:DUP 20226pri 20226red   (in the session header)

a=mid:20226pri                    (in the media section
                                of the primary 2022-6 flow)

a=mid:20226red                    (in the media section
                                of the secondary 2022-6 flow)
```

If SMPTE ST 2022-5 FEC packets are being transmitted by the sender, then an additional "a=group" clause shall be included in the session header, referencing the ST 2022-6 flow and the ST 2022-5 FEC packet flow, and the ST 2022-5 FEC flow shall have its own media section as follows:

```
a=group:FEC-FR 20226vid 20225vidfec   (in the session
header)

m= application <fecport> RTP/AVP 99 (media section for FEC)
c=IN IP4 233.252.0.1/127
a=rtpmap:99 SMPTE2022-5-FEC/27000000
a=fec-repair-flow: encoding-id=10  (TBD IANA id for ST
2022-5)
a=mid:20225vidfec                  (name ties to a=group
attr)
```

An example SDP object is shown in the appendix.

7. Studio-Profile Constraint Set

The following constraints taken together define a "Studio Profile" for uncompressed signals over IP networks, applicable in the context of an integrated production environment.

7.1. Well-Formed SDI

The SDI signals encapsulated within the St2022-6 flows shall be well-formed, meaning that they comply with the structures defined in the relevant SMPTE documents (ST 292, ST 274, ST 296, etc).

Note: RP-168 switching operations that are upstream of an IP encapsulator may cause one or more lines of an SDI video signal to contain incorrect number of samples (i.e. not in conformance with the relevant specification for the SDI signal) in the vicinity of the switching point. When such lines are encountered by an encapsulator, video sample data may need to be

added or removed from these lines prior the encapsulation process, thereby ensuring that every line within an ST-2022-6 encapsulated stream contains the same number of video samples.

7.2. Regarding Seamless Merge and FEC

When ST 2022-7 seamless merge is used, the specification of receiver class A (ST 2022-7 section 7) applies.

Devices compliant to this document may implement ST 2022-7 seamless protection and/or ST2022-5 FEC, but are not required to do so by this document. Receiving devices shall be tolerant of the presence of ST 2022-5 or ST 2022-7 datagrams if they are transmitted.

8. IANA Considerations (informative)

The VSF has asked SMPTE to petition the IANA for registration of the following MIME types relating to this document:

```
video/SMPTE2022-6

application/SMPTE2022-5-FEC
```

In addition, the VSF has asked SMPTE to petition the IANA for registration of FECFRAME ID number "10" (encoding-id=10) for SMPTE 2022-5 FEC.

The use of these IANA-managed types and IDs should be considered provisional until such time as their registration is completed.

Appendix A (Informative) Annotated Examples of SDP

This SDP object includes two SMPTE ST 2022-6 streams (main and protect via ST 2022-7), one AES67 audio stream with six channels, and one ANC data stream.

```
v=0
o=- 123456 11 IN IP4 192.168.1.1
s=Networked Media including 2022-6 video
i=A test of 2022-6 with AES67 via TR-03 plus TR-04
t=0 0
a=group:LS 20226pri A1 M1
a=group:DUP 20226pri 20226red
a=recvonly
m=video 51000 RTP/AVP 98 [note: primary video stream type ST 2022-6]
c=IN IP4 239.0.0.1/32
a=rtpmap:98 SMPTE2022-6/27000000
a=ts-refclk:ptp=IEEE1588-2008:39-A7-94-FF-FE-07-CB-D0:0
a=mediaclk:direct=2216659908
a=mid:20226pri

m=video 51100 RTP/AVP 98 [note: redundant video stream type ST 2022-6]
c=IN IP4 239.1.0.1/32
a=rtpmap:98 SMPTE2022-6/27000000
a=ts-refclk:ptp=IEEE1588-2008:39-A7-94-FF-FE-07-CB-D0:0
a=mediaclk:direct=2216659908
a=mid:20226red

m=audio 51200 RTP/AVP 97    [note: audio stream type AES67]
c=IN IP4 239.0.0.2/32
a=rtpmap:97 L24/48000/6
a=ptime:0.250
a=ts-refclk:ptp=IEEE1588-2008:39-A7-94-FF-FE-07-CB-D0:0
a=mediaclk:direct=963214424
a=fmtp:97 channel-order=SVIP.(L,C,R,Lrs,Rrs,LFE)
a=mid:A1

m=video 51300 RTP/AVP 98    [note: ANC draft-ietf-payload-rtp-ancillary ]
c=IN IP4 239.0.0.3/32
a=rtpmap:98 smpte291/90000
a=ts-refclk:ptp=IEEE1588-2008:39-A7-94-FF-FE-07-CB-D0:0
a=mediaclk:direct=2216659908
a=mid:M1
```