MISB MOTION IMAGERY STANDARDS BOARD	MISB RP 1302
RECOMMENDED PRACTICE	
Inserting KLV in Session Description Protocol (SDP)	**27 February 2014**

1 Scope

This MISB Recommended Practice (RP) presents a method to insert KLV (Key-Length-Value) encoded metadata in the Session Description Protocol (SDP) (RFC 4566). SDP is an Internet protocol format for describing streaming media initialization parameters, such as session announcement, session invitation, and parameter negotiation, in a multimedia session. SDP does not deliver media itself, but is used for negotiation between end-points of media type, format, and associated properties.

When systems that involve national security deliver motion imagery with KLV metadata over Real-time Transport Protocol (RTP), it is mandatory that the content be marked with a security classification and other security administration information to ensure proper handling. This RP provides a method for inserting such information into SDP.

Using this approach, session description and discovery is transmitted out-of-band from the streaming media data, thus conserving network bandwidth. Receivers only need request the SDP to determine if they are permitted to receive the stream, and may use the SDP to decide whether the stream is of interest before spending resources transmitting the content itself. SDP content size is substantially smaller than the streaming real-time data it is describing. Most SDP content is less than 5 Kilobytes. This makes for efficient use of network resources, where hundreds of simultaneous streaming multimedia sessions may be active.

2 References

2.1 Normative References

The following references and the references contained therein are normative.

[1] IETF RFC 4566 SDP: Session Description Protocol, Jul 2006
[2] SMPTE ST 336-2007 Data Encoding Protocol Using Key-Length-Value
[3] MISB ST 0102.10 Security Metadata Universal and Local Sets for Digital Motion Imagery, Oct 2013
[4] IETF RFC 4648 The Base16, Base32, and Base64 Data Encodings, Oct 2006
[5] MISB Motion Imagery Standards Profile (MISP) 6.6, Feb 2014

2.2 Informative References

[6] IETF RFC 3550 RTP: A Transport Protocol for Real-Time Applications, Jul 2003
[7] IETF RFC 3261 SIP: Session Initiation Protocol, Jun 2002
[8] IETF RFC 2326 Real Time Streaming Protocol (RTSP), Apr 1998
[9] IETF RFC 5234 Augmented BNF for Syntax Specifications: ABNF, Jan 2008

3 Terms and Definitions

End system

An application that generates content to be transmitted and/or consumes content to be received.

Multimedia session

One or more multimedia senders and receivers and the data streams flowing from senders to receivers. A multimedia conference is an example of a multimedia session.

RTP session

An association among a set of participants communicating with RTP. In a multimedia session, each media type is carried in a separate RTP stream. A participant distinguishes multiple RTP streams by using different pairs of destination transport addresses, where a pair of transport addresses comprises one network address plus a pair of ports for RTP and RTCP.

Session Description Protocol (SDP)

A text based format for describing streaming multimedia initialization parameters.

Transport address

The combination of a network address and port that identifies a transport-level endpoint; for example, an IP address and a UDP port. Packets are transmitted from a source transport address to a destination transport address.

4 Acronyms

IANA	Internet Assigned Numbers Authority
IP	Internet Protocol
RTCP	Real Time Control Protocol
RTP	Real Time Protocol
RTSP	Real Time Streaming Protocol
TS	Transport Stream
SDP	Session Description Protocol
SIP	Session Initiation Protocol
UDP	User Datagram Protocol

5 Revision History

Revision	Date	Summary of Changes
RP 1302	02/27/2014	• Initial Release

6 Introduction

The method outlined in this RP will only apply to systems announcing their multimedia session using SDP as part of a system feature for session discovery and/or classification.

Two methods are described:

- A process to insert KLV encoded metadata, such as security classification and security administration information, into the Session Description Protocol (SDP) format employing an SDP *attribute* field, which is the primary means for extending SDP. This RP mandates using the IANA registered attribute field `a=keywds:<value>` to insert a text representation of a KLV-encoded metadata set.
- A process for mapping KLV binary-based encoding to a text-based representation. Because SDP is a text-based protocol, this RP mandates the use of a base64 text representation, as outlined in IETF RFC 4648 [4], to encode binary-based KLV data.

The most likely employment of this RP is to mark security classification and security administration, or other identifying information of RTP session by adding an SDP attribute containing this information into the SDP session description. However, this RP is generic enough to allow systems that want to announce multimedia session using SDP.

For Internet Protocol (IP) based multimedia streaming, SDP is the dominant protocol employed to describe streaming media. Systems implementing this RP will only need to read the SDP session description to determine the security classification, or other identifying properties without parsing the streaming real-time data of a multimedia session.

Real-time Transport Protocol (RTP) (RFC 3550) [6] provides end-to-end transport functions for applications transmitting real-time data, such as audio, video or simulation data, over unicast or multicast IP networks. Applications typically use RTP over UDP to leverage RTPs multiplexing, packet count and checksum services. RTP is an alternative to MPEG-2 Transport Stream (TS) to stream multimedia content over an IP network.

Systems delivering video over MPEG-2 TS will typically contain security classification in the metadata elementary stream. MISB ST 0102 [3] describes security classification markings that are used in forming a motion imagery product.

Other protocols and mechanisms are required in addition to RTP to provide a usable service; in particular, session discovery, initiation and description. For end systems to participate in an RTP session, content providers typically distribute a session description; SDP (RFC 4566) [1] is the most often used protocol for that purpose. How a session description is disseminated is outside the scope of this RP. However, systems can employ Session Initiation Protocol (SIP) (RFC 3261) [7], Real-Time Streaming Protocol (RTSP) (RFC 2326) [8], or even email to distribute an SDP session description.

SDP provides a standard representation for information such as media details, transport addresses and other session description metadata to session participants. The SDP standard is independent of how this information is transported. SDP is purely a format for session description. An SDP session description includes the following:

- Session name and purpose
- Time(s) the session is active
- The media comprising the session
- Information needed to receive those media (addresses, ports, formats, etc.)

SDP is a structured text-based protocol, which can be machine processed and is human readable. An SDP session description is composed of a number of lines of text each in the form:

```
<type>=<value>
```

Where `<type>` must be exactly one case-significant character and `<value>` is structured text whose format depends on `<type>`.

A SDP session description consists of a session-level section followed by zero or more media-level sections. The session-level portion begins with a "`v=`" line and continues to the first media-level section. Each media-level section begins with an "`m=`" line and continues to the next media-level section, or end of the whole session description. In general, session-level values assume the default state for all media unless overridden by an equivalent media-level value. Below is an example of a SDP session description consisting of two RTP streams: one carrying audio and one carrying video.

```
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
```

Session Level Section

```
m=audio 49170 RTP/AVP 0
m=video 51372 RTP/AVP 99
a=rtpmap:99 h263-1998/90000
```

Media Level Section

Attributes are the primary means for extending SDP session description, which are in the form:

```
a=<attribute>
a=<attribute>:<value>
```

In the next section, a method to add security marking and classification information, or any other identifying properties using attributes and KLV-encoded metadata is discussed.

7 Method to insert KLV and security markings

KLV-encoded metadata is placed inside a SDP session description using the SDP attribute field `a=keywds:<KLV>`, which is an IANA registered value attribute defined in IETF RFC 4566 [1]. As described in IETF RFC 4566, publishers could insert this attribute field in a SDP session description so receivers can select a multimedia session based on keywords describing the purpose of the session. There is no central registry of keywords. Therefore, this RP permits the `keywds` value byte-string to contain both human readable text as well as text representation of KLV-encoded metadata. To facilitate parsing, keywords are space separated where a token `smpte336m=` is used to identify KLV data constructed using the practices of this RP.

Since SDP is a text-based protocol, a text representation of KLV-encoded metadata is required. IETF RFC 4648 [4] defines a base 64, base 32 and base 16 encoding scheme that is a commonly employed binary-to-text encoding standard. This RP mandates using base 64 encoding, where RFC 4648 refers this encoding as "base64".

Requirement	
RP 1302-01	The publisher end system shall convert a KLV binary sequence to a base64 text representation in accordance with IETF RFC 4648 [4].
RP 1302-02	The publisher shall insert the base64 text representation of a KLV metadata set into the SDP attribute `a=keywds`.
RP 1302-03	The publisher end system shall use Table 1, "The Base 64 Alphabet", in IETF RFC 4648 [4] for the valid encoding alphabet.
RP 1302-04	The publisher end system shall encode only metadata sets that conform to MISB MISP [5].
RP 1302-05	The subscriber end system shall decode base64 representations of a universal metadata set and a local metadata set in accordance with MISB MISP [5].
RP 1302-06	When base64 encoded KLV data is inserted into the value part of the attribute field `a=keywds`, the publisher end system shall prepend the token `smpte336m=` in front of the base64 encoded KLV data.
RP 1302-07	When human readable keywords and base64 encoded KLV data are inserted into the value part of the attribute field `a=keywds`, the publisher end system shall space separate keyword values.

A KLV-encoded metadata is inserted into the SDP attribute in the form:

```
a=keywds:smpte336m=<RFC 4648 Base64 encoded KLV data> CRLF
```

The following illustrates an example of a base64 text representation of a KLV-encoded metadata:

```
a=keywds:smpte336m=Bg4rNAILAQEOAQMBAQAAAIHsAggABH7EsYl3zjAqFRAGDis0AgsBAQ4BAw
EBAAAAAQEBAgECAwNDQU4GA0NBThQCAGYiAgQBGQIllBcELMUa5xgELn9GMDkEADQuBAwQR2VvY2V
udHJpYyBXR1M4NAsCRU8DCFhYWFhYWFhYGgL8ARwC/rAeAgN9IAIBJRsC/OcdAvqdHwICtCECBLMK
B1NreVN0YXI4AQAFAqFrBgKszQcCAAAEATIQAgREDQQszEFNDgQuhP0bEgQE6BtPEwT93d3eFAQAA
AAAEQIDMw8CGPEVBAAyX/tBAQIWAgUnSAgABH7EsYgJmAECNig=
```

If the publisher end-system wishes to include other human readable keywords in addition to KLV data, spaces are needed to delimit the keyword values. The following is an example where human readable keywords coexists with base64 encoded KLV data:

```
a=keywds: airfield smpte336m=<RFC 4686 Base64 encoded KLV data> foo bar CRLF
```

8 Keywds Value Grammar

This section provides an Augmented BNF (ABNF) grammar for the value part of an a=keywds:<keyword-values> attribute field. ABNF is defined in IETF RFC 5234 [9].

```
;keyword-values syntax for the value part of a=keywds:<keyword-values>

 keyword-values    =  keyword / keyword-values SP keyword CRLF
 keyword           =  text-keyword / klv-data-keyword
 klv-data-keyword  =  “smpte336m=” klv-data-value
 Klv-data-value    =  IETF RFC 4648 base64 encoded KLV data
 text-keyword      =  text
 text              =  byte-string
 byte-string       =  1*(%x01-09/%x0B-0C/%x0E-FF)
                      ;any byte except NUL, CR, or LF
;external references:
      ;CRLF, CR, LF, NUL, and SP are defined in IETF RFC 5234
```

9 Example – KLV metadata in a SDP session description

Assume a hypothetical multimedia streaming system that employs SDP to describe a multimedia session. To ensure proper security handling of the media stream and facilitate session discovery, a KLV set from the multimedia session is inserted into the SDP session description. The following is a binary sequence of a local data set (LDS) metadata set:

```
06 0E 2B 34 02 0B 01 01 0E 01 03 01 01 00 00 00 81 EC 02 08 00 04 7E C4 B1 89
77 CE 30 2A 15 10 06 0E 2B 34 02 0B 01 01 0E 01 03 01 01 00 00 00 01 01 01 02
01 02 03 03 43 41 4E 06 03 43 41 4E 14 02 00 66 22 02 04 01 19 02 0B 94 17 04
2C C5 1A E7 18 04 2E 7F 46 30 39 04 00 34 2E 04 0C 10 47 65 6F 63 65 6E 74 72
69 63 20 57 47 53 38 34 0B 02 45 4F 03 08 58 58 58 58 58 58 58 58 1A 02 FC 01
1C 02 FE B0 1E 02 03 7D 20 02 01 25 1B 02 FC E7 1D 02 FA 9D 1F 02 02 B4 21 02
04 B3 0A 07 53 6B 79 53 74 61 72 38 01 00 05 02 A1 6B 06 02 AC CD 07 02 00 00
04 01 32 10 02 04 44 0D 04 2C CC 41 4D 0E 04 2E 84 FD 1B 12 04 04 E8 1B 4F 13
04 FD DD DD DE 14 04 00 00 00 00 11 02 03 33 0F 02 18 F1 15 04 00 32 5F FB 41
01 02 16 02 05 27 48 08 00 04 7E C4 B1 88 09 98 01 02 36 28
```

This binary sequence is converted to base64 to create a base64 text representation, which looks as follows:

```
Bg4rNAILAQEOAQMBAQAAAIHsAggABH7EsYl3zjAqFRAGDis0AgsBAQ4BAwEBAAAAAQEBAgECAwNDQ
U4GA0NBThQCAGYiAgQBGQIllBcELMUa5xgELn9GMDkEADQuBAwQR2VvY2VudHJpYyBXR1M4NAsCRU
```

```
8DCFhYWFhYWFhYGgL8ARwC/rAeAgN9IAIBJRsC/OcdAvqdHwICtCECBLMKB1NreVN0YXI4AQAFAqF
rBgKszQcCAAAEATIQAgREDQQszEFNDgQuhP0bEgQE6BtPEwT93d3eFAQAAAAAEQIDMw8CGPEVBAAy
X/tBAQIWAgUnSAgABH7EsYgJmAECNig=
```

This text string is then inserted into an SDP session description using the SDP attribute `a=keywds` as shown below:

```
v=0
o=- 15075599483178459290 15075599483178459290 IN IP4 nobody-PC
s=Unnamed
i=N/A
c=IN IP4 239.3.3.71/255
t=2208988800 0
a=tool:vlc 1.1.7
a=keywds:smpte336m=Bg4rNAILAQEOAQMBAQAAAIHsAggABH7EsYl3zjAqFRAGDis0AgsBAQ4BAw
EBAAAAAQEBAgECAwNDQU4GA0NBThQCAGYiAgQBGQIllBcELMUa5xgELn9GMDkEADQuBAwQR2VvY2V
udHJpYyBXR1M4NAsCRU8DCFhYWFhYWFhYGgL8ARwC/rAeAgN9IAIBJRsC/OcdAvqdHwICtCECBLMK
B1NreVN0YXI4AQAFAqFrBgKszQcCAAAEATIQAgREDQQszEFNDgQuhP0bEgQE6BtPEwT93d3eFAQAA
AAAEQIDMw8CGPEVBAAyX/tBAQIWAgUnSAgABH7EsYgJmAECNig=
a=recvonly
a=type:broadcast
a=charset:UTF-8
m=video 50000 RTP/AVP 96
b=AS:1024
b=RR:0
a=rtpmap:96 H264/90000
a=fmtp:96 packetization-mode=1;profile-level-id=64001e;sprop-parameter-
sets=Z2QAHqzZQLQSbARAAAADAEAAAAyjxYtlgA==,aOvssiw=;
m=application 50000 RTP/AVP 97
c=IN IP4 239.3.2.71/255
a=rtpmap:97 smpte336m/4
```

A subscriber reads the above SDP and decodes the attribute field value for `a=keywds`, which returns the KLV metadata to its original binary format.