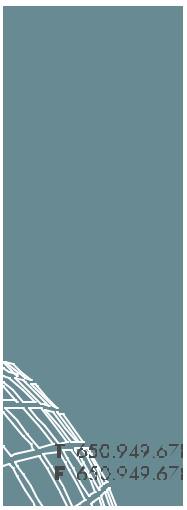
ISMA SPECIFICATION DISCLAIMER

"LEGAL DISCLAIMER

THIS SPECIFICATION IS PROVIDED "AS IS" AND WITHOUT ANY WARRANTY OR REPRESENTATION OF ANY KIND, EXPRESS OR IMPLIED. WITHOUT LIMITATION, THERE IS NO WARRANTY OF NON-INFRINGEMENT, NO WARRANTY OF MERCHANTABILITY, AND NO WARRANTY OF FITNESS FOR A PARTICULAR PURPOSE. ALL WARRANTIES ARE EXPRESSLY DISCLAIMED.

USER ASSUMES THE FULL RISK OF USING THE SPECIFICATION. IN NO EVENT SHALL ISMA OR ANY MEMBER OF ISMA BE LIABLE FOR ANY ACTUAL, DIRECT, INDIRECT, PUNITIVE, OR CONSEQUENTIAL DAMAGES ARISING FROM SUCH USE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGES."



2570 W. El Camino Real, Suite 304
Mountain View, CA 94040

Internet Streaming Media Alliance Implementation Specification

Version 1.0 + Corrigenda

June 3, 2004

Internet Streaming Media Alliance
www.isma.tv
Presidio of San Francisco
P.O.Box 29920
572B Ruger Street
San Francisco, CA 94129-0920
Phone: (+1) 415. 561.6276
Fax: (+1) 415. 561. 6120

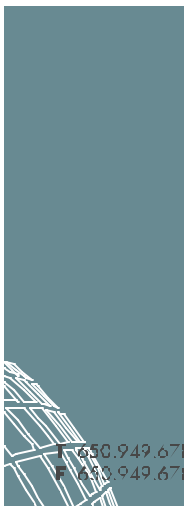
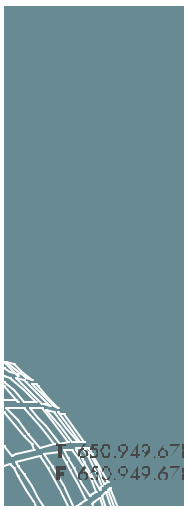
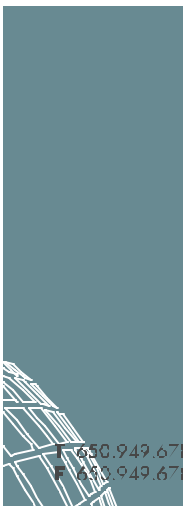


Table Of Contents



1 Acronyms & Terms

AVP	Audio Visual Profile (IETF RFC 3551)
BIFS	Binary Format for Scene
CIF	Common Intermediate Format (352 x 288)
cRTP	Compressed Real Time Protocol (IETF RFC 2508)
ESD	Elementary Stream Descriptor
IEC	International Electrotechnical Commission
IETF	Internet Engineering Task Force
IOD	Initial Object Descriptor
IP	Internet Protocol
IPv4	Internet Protocol Version 4
IPv6	Internet Protocol Version 6
ISMA	Internet Streaming Media Alliance
ISO	International Organization for Standardization
MP4	MPEG-4 File Format
M4IF	MPEG-4 Industry Forum
OCR	Object Clock Reference
OD	Object Descriptor
QCIF	Quarter Common Intermediate Format (176 x 144)
QoS	Quality of Service
RFC	Request for Comment
RTP	Real Time Protocol (IETF RFC 3550)
RTSP	Real Time Streaming Protocol (IETF RFC 2326)
SDP	Session Description Protocol (IETF RFC 2327)
TCP	Transmission Control Protocol (IETF RFC 793)
UDP	User Datagram Protocol (IETF RFC 768)
WMF	Wireless Multimedia Forum



Internet Streaming Media Alliance Implementation Specification

2 Technical Specification:

2.1 Document Status

This version of the document is a proposed final version from the Technical committee to the ISMA Board of Directors for approval.

2.2 Scope

For several years, the promise of delivering video and audio over the Internet has been widely promoted. However, the fulfillment of these promises has been delayed due to a number of technical issues. One significant impediment has been the lack of a set of widely adopted technical standards for the transmission of video and audio on the Internet. This has led to a fragmentation of the market and a lack of critical mass necessary to achieve rapid deployment.

This implementation specification addresses this need by setting forth a framework for the use of existing open standards that vendors can use to build interoperable video and audio systems for use on IP networks and the Internet.

Note that this specification assumes use of the existing IPv4 infrastructure and does not require any deployments of new or advanced IP technology. However, this specification will allow conforming implementations to benefit from improvements in the IP infrastructure as they are deployed. Specifically, this implementation specification is independent of the existence of IPv6, IPsec, IP multicast, and IP QoS.

Note that this specification targets broadband quality networks (less than 1.5Mbps in raw bandwidth). The 1.0 Implementation Specification also addresses both narrowband (dialup) and wireless networks.
Further this specification assumes use of existing MPEG technologies and although initially focused on MPEG-4 technologies, future adaptations and revisions may include MPEG-2, 7 and other non-MPEG technologies.

2.3 Architecture

At its simplest, the architecture of ISMA consists of a media server, an IP network, and a media client. Media is transmitted from the media server to the media client in one of two modes: on-demand or broadcast.

In the on-demand mode, the media client requests media content from the media server that after appropriate verifications transmits the media to the client.
In the broadcast mode, the media server begins transmitting the media content to one or more media clients at a specified time. The media client is advised when to receive the media content via an out-of-band mechanism.
This simple model can be easily extended to allow for a number of intermediary systems in the transmission process that perform a range of services. These intermediary systems either work transparently from the viewpoint of both the media server and media client, or the intermediary system presents a media server interface to media clients, and a media client interface to media servers. Examples of the services that can

be provided by intermediary systems are: short-term or long-term storage, re-encoding to new bit rates, mixing with other media streams. Streaming media caches/proxies are the current preeminent example of such intermediary systems.

Note this architecture makes no assumptions that the transmission of media across the IP network must occur in “real-time” such that the media can be immediately rendered at the media client. In fact, we wish to enable media to flow in a mixture of real-time, and off-line transmissions, so as to best utilize network resources, and provide the best end-user experience.

Also note that this architecture makes no assumptions regarding the “production” phase of media content where capture, editing, and encoding of the media occurs. How such systems operate and what output they produce is an implementation issue for vendors producing media servers, encoders and the like.

2.4 Technology Selection Criteria

Three criteria are used in selecting the technologies for this implementation specification:

The first criterion is that the technology must be an open standard. The specifications must be available to all who wish to participate and the licensing of any associated intellectual property should be on reasonable terms for vendors of any size.

The second criterion is that the technology should enable interoperable products to be brought to market quickly. This argues for giving preference to technologies that already have multiple implementations and have wide deployment. This criterion also argues that the profile should aim for simplicity; seeking a solution that addresses the core needs of the Internet video and audio applications, and not the specialized needs of every conceivable application.

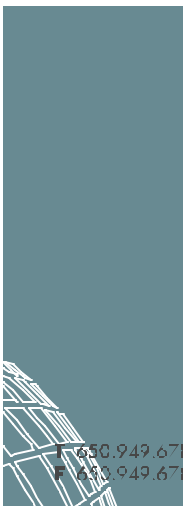
The third criterion is that the technology be forward-looking, and provide for new network media technologies and emerging video-enabled information appliances.

Note that these criteria may evolve in future versions of this document as more sophisticated standards and implementations evolve in the marketplace.

2.5 Definitions

Note this document uses the following definitions taken from IETF RFC 2119:

1. MUST This word, or the terms "REQUIRED" or "SHALL", mean that the definition is an absolute requirement of the specification.
2. SHOULD This word, or the adjective "RECOMMENDED", mean that there may exist valid reasons in particular circumstances to ignore a particular item, but the full implications must be understood and carefully weighed before choosing a different course.
3. MAY This word, or the adjective "OPTIONAL", mean that an item is truly optional. One vendor may choose to include the item because a particular marketplace requires it or because the vendor feels that it enhances the product while another vendor may omit the same item. An implementation that does not include a particular option MUST be prepared to interoperate with another implementation that does include the option, though perhaps with reduced functionality. In the same vein an



implementation that does include a particular option MUST be prepared to interoperate with another implementation that does not include the option (except, of course, for the feature the option provides.)

2.6 Functions

The functions of the Internet video and audio ecosystem can be described as follows.

2.6.1 Media Transmission

This is the means by which the media content is transmitted over the network, either in real-time or off-line.

2.6.2 Media Control

This is the means by which a viewer requests transmission of media content to them for rendering and optionally controls the transmission (e.g. pause, rewind, fast-forward).

2.6.3 Media Announcement

This is the means by which a viewer (either human or computer) discovers the existence of available media content and the information necessary to request access to the media content.

Note that a given single product may not implement all of these functions, but will contribute to the ecosystem's functionality in one or more areas. For example, a specialized video encoder product might only address the media transmission function.

2.7 Profiles

The following technical standards are the identified profiles for the ecosystem. To be compliant with this specification a product must completely implement profile 0, and may implement additional profiles. For example, an on-demand video server would likely need to implement all possible profiles to address a wide client base. However a decoder/terminal may only support profile 0.

This approach has been taken to ensure that any product certified as ISMA compliant, has the capability to minimally interoperate with any other ISMA compliant product.

This is a definition of base interoperability. Vendors are still free to add additional functionality beyond that specified in this document. However that said, a conforming product cannot make any additional requirements beyond this specification to interoperate with another conforming system.

2.7.1 All Profile Requirements

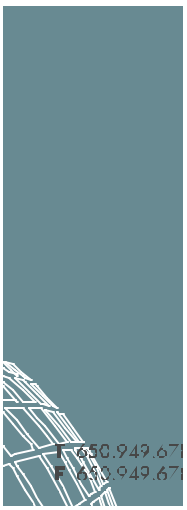
The following lists those requirements common to all profiles.

Media Decoding

A conforming implementation MUST implement either Video, or Audio, or both.

Transports

REQUIRED - RTP: A Transport Protocol for Real-Time Applications RTP
IETF RFC 3550



REQUIRED - RTP Profile for Audio and Video Conferences with Minimal Control
IETF RFC 3551
REQUIRED – User Datagram Protocol
IETF RFC 768

OPTIONAL - Interleaved RTSP & RTP/AVP over TCP transport
Real Time Streaming Protocol – IETF RFC 2326, section 10.12
Transmission Control Protocol - IETF RFC 793

Rationale: The RTP/AVP/UDP profile is the simplest and most widely supported option in current Internet streaming media systems. The RTSP/RTP interleave over TCP provides the option of reliable transport. Further RTSP/RTP over TCP will also permit traversal of Network Address Translators and Firewalls.

The media may be transmitted in unicast, multicast or in a broadcast manner. For the transmission of numerous simultaneous sessions it is RECOMMEDED that multicast be used for media transmission. It is further RECOMMENDED that a multicast implementation be conformant to IGMP v3. See section 4 for IGMPv3 references at time of publication.

RTP Payloads
REQUIRED (Video) - RFC 3016 RTP Payload Format for MPEG-4 Audio/Visual Streams
REQUIRED (Audio) - RFC 3460 RTP Payload Format for MPEG-4 Streams

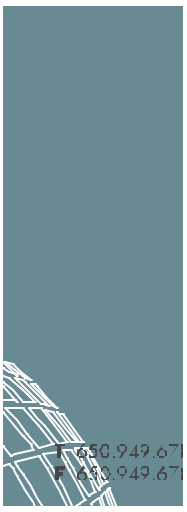
ISMA Restriction: Video decoder specific configuration information MUST be present in the SDP description of the media stream. This is a restriction of RFC 3016 that requires it to be in-band, at least once, and optionally present in the SDP description.

Note: cRTP – IETF RFC 2508 is an optional adjunct to RTP and can be used in conjunction with this specification in a transparent fashion.

Content Distribution
REQUIRED – ISO base media file format - ISO/IEC 14496-12:2003 | 15444-12:2003
REQUIRED – MP4 file format - ISO/IEC 14496-12:2003

Rationale – In order to encourage content distribution and interoperability between products at the file storage level, the MPEG-4 MP4 Format, based on Apple's QuickTime file format, is the most obvious choice for MPEG-4 content.

ISMA compliant content, when stored in files, will contain the minimal BIFS and OD streams as described in Appendix E. The BIFS and OD streams MUST be stored in their own tracks. The usage of "data:" URLs to carry these streams (as described in Appendix F) is limited to transmission sessions. Of course, files may contain hint tracks to assist the preparation of the IOD for transmission. This restriction is required in order to avoid multiple flavors of same content.
ISMA compliant content shall also adhere to the following limitations to the ISO base media file format:



- compact sample sizes ('stz2') shall not be used;
- padding bits boxes (‘padb’) shall not be used;
- movie fragments shall not be used;
- In the Data Reference atom of MP4 files which are to be placed on servers (hinted or un-hinted), the data sources are restricted to being from the following set:
    - the same file as the movie atom meta-data ('self-contained')
    - a relative URL (interpreted as relative to the file containing the movie atom meta-data)
    - an absolute URL using the "file:" 'protocol'.

Note also that ISO/IEC 14496-14, section 5.6, constraints the number of sample entries to be one for video, audio, bifs and OD tracks.

RECOMMENDED - MPEG MP4 Format:

In order to provide easy access to the sample dimensions, the initial 32 bit reserved field in VisualSampleEntry should be used as a 16 bits width followed by a 16 bit height of the video frame.

To facilitate a simple parsing scheme, the atoms that provide information for other atoms should be placed before the dependent atoms. For example, the header atom for a container should precede other atoms in the container that require information found in the header atom, the handler reference atom should precede the relevant information atom, and the Decoding Time to Sample Atom should precede the Composition Time to Sample Atom.

Media Control

REQUIRED - Real Time Streaming Protocol (RTSP)
IETF RFC 2326

The minimal RTSP implementation described in Appendix D of RFC2326 SHALL be used, with the addition of the DESCRIBE method.

RECOMMENDED – that a conforming RTSP implementation support the DESCIBE method. If the DESCRIBE method is used, it is REQUIRED that SDP is supported as the description format, as specified in Appendix C of RFC2326.

REQUIRED – since the RTP/AVP transport must be supported to conform to this specification. A conforming RTSP implementation MUST support RTP/AVP in the “Transport” header and MUST support the “RTP-Info” header. It SHOULD also support the “client_port”, “server_port”, "source", and “ssrc” parameters of the “Transport” header.

REQUIRED – that a conforming RTSP implementation support the PLAY method.

RECOMMENDED – that a conforming RTSP implementation accept/generate the RTSP headers: “Bandwidth”, “Cache-control”, “Expires”, “If-modified-since”, “Last-modified”, “User-Agent”, “Server”.

Media Announcement
REQUIRED - SDP: Session Description Protocol
IETF RFC 2327

The SDP data should be formatted according to SDP specification [RFC2327] and Appendix C of RFC2326.

Use the ISMA specific parameter:

a=isma-compliance:<profile>,<lowest-spec-version>,<authored-to-version>
where:
profile:
an integer specifying the ISMA profile to which the content conforms
lowest-spec-version:
a decimal number, indicating the lowest version number of the ISMA specification to which a client can conform, and still decode the content. Clients MUST not decode content with a lowest-spec-version higher than the highest specification version that they implement. The first published specification was 1.0; therefore this field is 1.0 or greater
authored-to-version:
the version of the specification against which the content was authored. Ideally the client also implements this version, whereupon the user can be more confident that the content is being completely decoded. A content author may choose to allow clients written to earlier versions of the spec achieve partial decode.

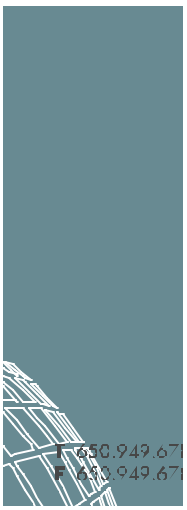
Servers MUST generate the SDP iod attribute using a binary iod (not XML); see Appendix E for details. If the URL is a data URL, then the binary IOD is encapsulated directly in the URL using base64 encoding.

Note that the media description in SDP form can be transmitted in a number of ways; examples include HTTP, RTSP, SMTP, SAP, and SIP.

2.7.2 Profile 0

Rationale: This profile was selected to allow for video and audio at bitrates suitable that match capabilities of narrowband and mobile wireless infrastructures and to align with the patent pool work in M4IF.

Video
REQUIRED - MPEG-4 ISO/IEC 14496-2:2001
Simple Profile @ Level 1
Typical Visual Session Size is QCIF (176x144)
Maximum bit rate is 64kbit/s



ISMA Restriction: Profile 0 is limited to one (1) video object only

Audio
REQUIRED - MPEG-4 ISO/IEC 14496-3:2001
High Quality Audio Profile @ Level 2
Up to 2 channels
Up to 48000 Hz sampling rate
This profile contains both CELP and Low Complexity AAC

ISMA Restriction: Profile 0 is limited to one (1) audio object only

Informative Note: Using the High Quality Audio Profile @ L2 may lead to a situation where the audio bit rate is greater than the video bit rate. Although unusual, there are scenarios where this may be desired.

2.7.3 Profile 1

Rationale: This profile was selected to allow for a richer streaming experience over infrastructures with broadband bit rates. Profile 1 is a superset of Profile 0 and can fully decode any streams generated by Profile 0 encoders.

Video
REQUIRED - MPEG-4 ISO/IEC 14496-2:2001 + Amd 2:2001
Advanced Simple Profile @ Level 3
Typical Visual Session Size is CIF (352x288)
ISMA Maximum Bitrate 1.5 Mbps

ISMA Restriction: Profile 1 is limited to one (1) video object only

Audio
REQUIRED - MPEG-4 ISO/IEC 14496-3:2001
High Quality Audio Profile @ Level 2
Up to 2 channels
Up to 48000 Hz sampling rate
This profile contains both CELP and Low Complexity AAC

ISMA Restriction: Profile 1 is limited to one (1) audio object only

Cumulative Bit Rate Limitation
ISMA Restriction: The combined audio and video bitrates in a Profile 1 session is limited to 1.5 Mbps.

3 Security Considerations

The current version of this framework does not add nor does it detract from any security that exists with the protocols/specifications it uses/refers to.

Future versions of this framework must address both network level security issues as well as digital rights management issues.

4 Future Work Areas

The following areas are possibilities for inclusion in a future version of this document.

Digital Rights Management:

MPEG-4 IPMP provides a framework for digital rights management (DRM), but today there is no standard that addresses the complete function of a DRM system. The ISO MPEG-4 group is currently soliciting submissions to fully address DRM. We would expect to consider the results of this standards effort for inclusion in ISMA.

Additional MPEG-4 Profiles:

Those that provide scalable encoding (aka temporal or spatial enhancement levels), and a mapping of the encoding levels to RTP streams such that network elements can provide differential QoS and/or rate-adaptation.

Profiles targeted at home entertainment level qualities, i.e. Main Profile @ L3, and Main Profile @ L4 or ACE profile.

Additional MPEG-4 Technologies: MPEG-4 Systems, or Synthetic Video, or Synthetic Audio.

Investigation into various techniques for firewall traversal. These could include using HTTP or RTSP proxies.

An RTSP mechanism for negotiating QoS for video and audio streams – there exist some proprietary schemes today, but no standard yet exists.

An RTP/RTCP retransmission and/or forward-error correction (FEC) mechanism – several proposals have been made within the IETF but it is still early in the standardization process for these capabilities.

Network level Quality of Service.

5 References

Internet Engineering Task Force: www.ietf.org
Motion Picture Experts Group http://www.cselt.it/mpeg/

Postel, J., "User Datagram Protocol," RFC 768, USC/Information Sciences Institute, August 1980.

Postel, J, “Transmission Control Protocol”, RFC 793, USC/Information Sciences Institute, September 1981.

H. Schulzrinne, S. Casner, R. Frederick, and V. Jacobson, "RTP: A transport protocol for real-time applications." RFC 3550, July 2003.

H. Schulzrinne, S. Casner, “RTP Profile for Audio and Video Conferences with Minimal Control”, RFC 3551, July 2003.

S. Casner, V. Jacobson, “Compressing IP/UDP/RTP Headers for Low Speed Serial Links”, RFC 2508, February 1999.

H. Schulzrinne, A. Rao, R. Lamphier, “Real Time Streaming Protocol (RTSP)”, RFC 2326, April 1998

M. Handley, V. Jacobson, “SDP: Session Description Protocol”, RFC 2327, April 1998.

Y. Kikuchi, T. Nomura, S. Fukunaga, Y. Mutsui, H.Kimata, “RTP Payload Format for MPEG-4 Audio/Visual Streams”,RFC 3016, November 2000.

J. Van Der Meer et al., "RTP Payload Payload Format for MPEG-4 Streams." RFC 3640, Nov 2003.

The following are MPEG references listed within this document:

ISO/IEC JTC1/SC 29/WG11 ISO/IEC 14496-1:2003, “Information technology – Coding of audio-visual objects – Part 1: Systems”

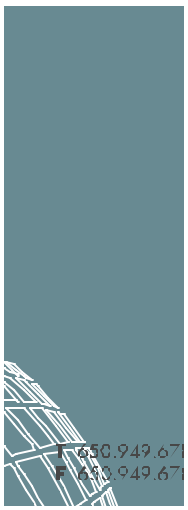
ISO/IEC JTC1/SC 29/WG11 ISO/IEC 14496-2:2001, “Information technology – Coding of audio-visual objects – Part 2: Visual”

ISO/IEC JTC1/SC 29/WG11 ISO/IEC 14496-2:2001-AMD2, “Information technology – Coding of audio-visual objects – Part 2: Visual, Amendment 2: Streaming video profile”

ISO/IEC JTC1/SC 29/WG11 ISO/IEC 14496-3:2001, “Information technology – Coding of audio-visual objects – Part 3: Audio”

ISO/IEC JTC1/SC 29/WG11 ISO/IEC 14496-8, “Information technology – Coding of audio-visual objects – Part 8: Carriage of ISO/IEC 14496 contents over IP networks”

ISO/IEC JTC1/SC 29/WG11 ISO/IEC 14496-12:2003 | 15444-12:2003, “Information technology – Coding of audio-visual objects – Part 12: ISO base media file format” &

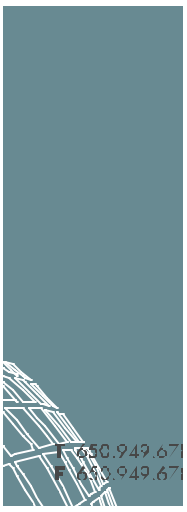


"Information technology – JPEG 2000 image coding system – Part 12: ISO base media file format"

ISO/IEC JTC1/SC 29/WG11 ISO/IEC 14496-14:2003, "Information technology – Coding of audio-visual objects – Part 14: MP4 file format"

6 Modification History

Version	Date	Editor	Changes
1.0.	August 27, 2001	Art Howarth	First public specification.
1.0.1	June 3, 2004	Guido Franceschini Carolyn Malestic	Integration of ISMA 1.0 and corrigenda, reference changes, other editorial changes

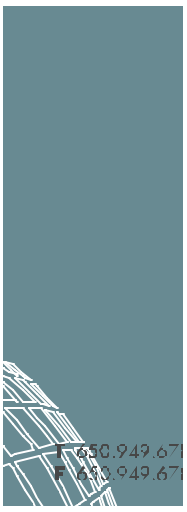


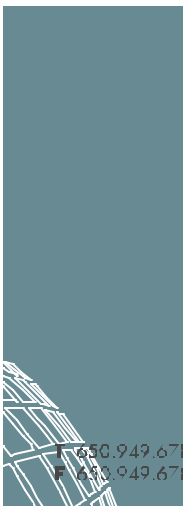
7 Contacts

ISMA Web Site

www.isma.tv

Implementation Specific Editor	SpecEditor@isma.tv
Executive Director of ISMA	ExecDirector@isma.tv





8 Appendix A: RTP/RTCP, RTSP, SDP Field Usage (Normative)

This appendix lists all parameters in RTP/RTCP, RTSP and SDP specifications that have been listed as optional that for ISMA purposes need to be mandatory, or fixed to certain values.

RTP/RTCP:

No parameters to date have been noted as needing to be fixed or mandatory.

RTSP:

No parameters to date have been noted as needing to be fixed or mandatory.

SDP:

a=isma-compliance:<profile>,<lowest-spec-version>,<authored-to-version>

where:

profile:

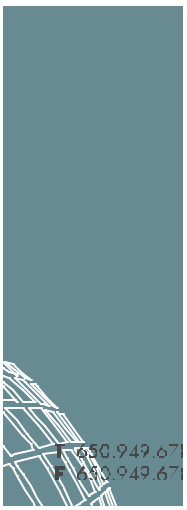
an integer specifying the ISMA profile to which the content conforms

lowest-spec-version:

a decimal number, indicating the lowest version number of the ISMA specification to which a client can conform, and still decode the content. Clients MUST not decode content with a lowest-spec-version higher than the highest specification version that they implement. The first published specification was 1.0; therefore this field is 1.0 or greater

authored-to-version:

the version of the specification against which the content was authored. Ideally the client also implements this version, whereupon the user can be more confident that the content is being completely decoded. A content author may choose to allow clients written to earlier versions of the spec achieve partial decode.

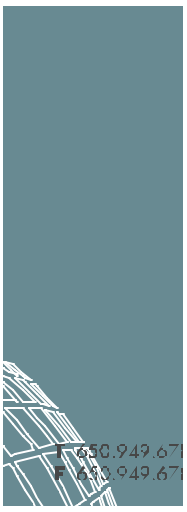


9 Appendix C: Hint Track Example (Informative)

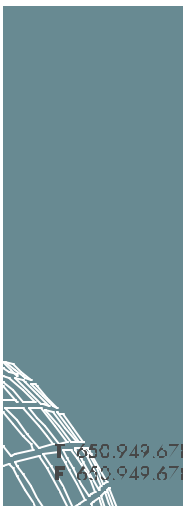
The following is an example of a hint track with 3 samples. Six packets are constructed from these three samples

```
mpeg4video.mov {
  time scale:                                   600
  duration:                                     60
  Descriptions {
    'rtp ' (14 octets):
      sdp b=AS:159
            73 64 70 20 62 3D 41 53 3A 31 35 39 0D 0A
  }
  Hint Tracks {
    Hinted Video Track {
      duration:                                 60
      media time scale:                         600
      Hinted Tracks {
        1:                                      'vide' track ID=1
      }
      Hint Info {
        payload-type:                           96 (X-ISMA)
        Descriptions {
          'sdp ' (73 octets):
           m=video 0 RTP/AVP 96
           b=AS:159
           a=rtpmap:96 X-ISMA
           a=control:trackID=3
        }
      }
      Sample Descriptions {
        Sample Description 1 {
          size:                                 36
          format:                               'rtp '
          data reference index:                 1
          hint track version:                   1
          highest compatible version:           1
          max packet size:                      1450
          Data {
            'tims' (4 octets):
              ....                  00 01 5F 90
          }
        }
      }
      Samples (15) {
        Sample 1 {
          sample time:                          0
          sample description index:             1
          Packets (4) {
            Packet 0 {
              RTP sequence number seed:         1
              relative transmission time:       0
              repeat-flag:                      false
              B-frame-flag:                     false
              RTP Header {
                P-bit:                          false
                X-bit:                          false
```

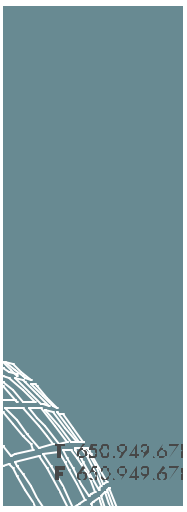
```
    M-bit:                          false
    payload-type:                   96
  }
  Payload (1 data entries) {
    data entry 0, constructor type 2 (sample data):
    track reference:                -1 (Hinted Video Track)
    sample number:                  1
    sample offset:                  116
    length:                         1438
    bytes per block:                1
    samples per block:              1
  }
}

Packet 1 {
  RTP sequence number seed:         2
  relative transmission time:       0
  repeat-flag:                      false
  B-frame-flag:                     false
  RTP Header {
    P-bit:                          false
    X-bit:                          false
    M-bit:                          false
    payload-type:                   96
  }
  Payload (1 data entries) {
    data entry 0, constructor type 2 (sample data):
    track reference:                -1 (Hinted Video Track)
    sample number:                  1
    sample offset:                  1554
    length:                         1438
    bytes per block:                1
    samples per block:              1
  }
}

Packet 2 {
  RTP sequence number seed:         3
  relative transmission time:       0
  repeat-flag:                      false
  B-frame-flag:                     false
  RTP Header {
    P-bit:                          false
    X-bit:                          false
    M-bit:                          false
    payload-type:                   96
  }
  Payload (1 data entries) {
    data entry 0, constructor type 2 (sample data):
    track reference:                -1 (Hinted Video Track)
    sample number:                  1
    sample offset:                  2992
    length:                         1438
    bytes per block:                1
    samples per block:              1
  }
}
```

```
            Packet 3 {
              RTP sequence number seed:         4
              relative transmission time: 0
              repeat-flag:                      false
              B-frame-flag:                     false
              RTP Header {
                P-bit:                          false
                X-bit:                          false
                M-bit:                          true
                payload-type:                   96
              }
              Payload (1 data entries) {
                data entry 0, constructor type 2 (sample data):
                track reference:                -1 (Hinted Video Track)
                sample number:                  1
                sample offset:                  4430
                length:                         350
                bytes per block:                1
                samples per block:              1
              }
            }
          }
        }
Sample 2 {
          sample time:                          20
          sample description index:             1
          Packets (1) {
            Packet 0 {
              RTP sequence number seed:         5
              relative transmission time:       0
              repeat-flag:                      false
              B-frame-flag:                     false
              RTP Header {
                P-bit:                          false
                X-bit:                          false
                M-bit:                          true
                payload-type:                   96
              }
              Payload (1 data entries) {
                data entry 0, constructor type 2 (sample data):
                track reference:                -1 (Hinted Video Track)
                sample number:                  2
                sample offset:                  32
                length:                         428
                bytes per block:                1
                samples per block:              1

              }
            }
          }
        }
Sample 3 {
          sample time:                          40
          sample description index:             1
          Packets (1) {
            Packet 0 {
              RTP sequence number seed:         6
```

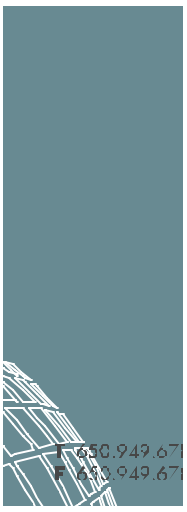
```
              relative transmission time:       0
              repeat-flag:                      false
              B-frame-flag:                     false
              RTP Header {
                P-bit:                          false
                X-bit:                          false
                M-bit:                          true
                payload-type:                   96
              }
              Payload (1 data entries) {
                data entry 0, constructor type 2 (sample data):
                track reference:                -1 (Hinted Video Track)
                sample number:                  3
                sample offset:                  32
                length:                         364
                bytes per block:                1
                samples per block:              1
              }
            }
          }
        }
      }
    }
  }
}
```

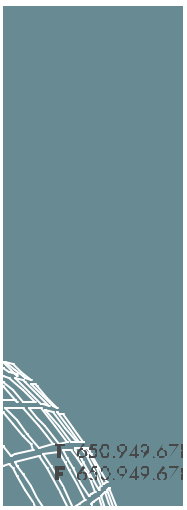
10 Appendix D: RTSP and SDP Examples (Informative)

The following is an example of an RTSP session.

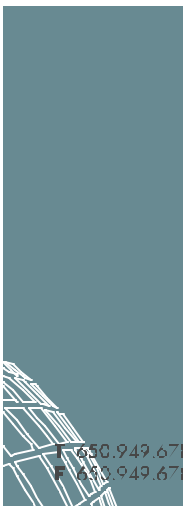
```
DESCRIBE rtsp://stream.isma.tv/foo.mp4 RTSP/1.0
CSeq: 1
Accept: application/sdp
Bandwidth: 4294967286
Accept-Language: en-US
User-Agent: ISMA-Client


RTSP/1.0 200 OK
Server: ISMA-Server
Cseq: 1
Content-length: 853
Content-Type: application/sdp
Content-Base: rtsp://stream.isma.tv/foo.mp4/

v=0
o=- 2890844256 2890842807 IN IP4 17.202.35.74
s=/foo.mp4
t=0 0
u=http://isma.tv/
e=admin@isma.tv
c=IN IP4 0.0.0.0
a=control:/
a=range:npt=0-   7.20000
a=ISMA-compliance:1,1.0,1
a=mpeg4-iod: "data:application/mpeg4-iod;base64,
AoF/AE8BAQEBAQOBEgABQHRkYXRh
OmFwcGxpY2F0aW9uL21wZWc0LW9kLWF1O2Jhc2U2NCxBVGdCR3dVZkF4Y0F5U1FBWlFRTklCRUFG
M0FBQVBvQUFBRERVQVlCQkFFWkFwOERGUUJsQlFRTlFCVUFCOUFBQUQ2QUFBQStnQVlCQXc9PQQN
/wUAAMgAAAAAAAAAAYJAQAAAAAAAAAAA2EAAkA+ZGF0YTphcHBsaWNhdGlvbi9tcGVnNC1iaWZz
LWF1O2Jhc2U2NCx3QkFTZ1RBcUJYSmhCSWhRUlFVL0FBPT0EEgINAAAUAAAAAAAAAAAFAwAAQAYJ
AQAAAAAAAAAA"

m=audio 0 RTP/AVP 96
a=rtpmap:96 mpeg4-generic/44100/2
a=control:trackID=5
a=fmtp:96 streamtype=5; profile-level-id=15; mode=AAC-hbr;
config=9122620000; sizelength=13; indexlength=3;
indexdeltalength=3;
a=mpeg4-esid:101
m=video 0 RTP/AVP 97
a=rtpmap:97 MP4V-ES
a=control:trackID=6
a=fmtp:97 profile-level-id=1;
config=000001B001000001B5090000010000000120008440FA282C2090A21F
a=mpeg4-esid:201

SETUP rtsp://stream.isma.tv/foo.mp4/trackID=6 RTSP/1.0
CSeq: 2
```

```
Transport: RTP/AVP;unicast;client_port=6970-6971
User-Agent: ISMA-Client
Accept-Language: en-US


RTSP/1.0 200 OK
Server: ISMA-Server
Cseq: 2
Session: 2160059768067155308;timeout=60
Transport: RTP/AVP;unicast;client_port=6970-
6971;source=17.202.35.74;server_port=2000-2001


SETUP rtsp://stream.isma.tv/foo.mp4/trackID=5 RTSP/1.0
CSeq: 3
Transport: RTP/AVP;unicast;client_port=6972-6973
Session: 2160059768067155308
User-Agent: ISMA-Client
Accept-Language: en-US


RTSP/1.0 200 OK
Server: ISMA-Server
Cseq: 3
Session: 2160059768067155308
Transport: RTP/AVP;unicast;client_port=6972-
6973;source=17.202.35.74;server_port=2000-2001


PLAY rtsp://stream.isma.tv/foo.mp4 RTSP/1.0
CSeq: 4
Range: npt=0.000000-7.200000
Session: 2160059768067155308
User-Agent: ISMA-Client


RTSP/1.0 200 OK
Server: ISMA-Server
Cseq: 4
Session: 2160059768067155308
RTP-Info: url=trackID=6;seq=18230;rtptime=1426696138
 ,url=trackID=5;seq=27737;rtptime=1269965233


PAUSE rtsp://stream.isma.tv/foo.mp4 RTSP/1.0
CSeq: 5
Session: 2160059768067155308
User-Agent: ISMA-Client


RTSP/1.0 200 OK
Server: ISMA-Server
```

```
Cseq: 5
Session: 2160059768067155308


PLAY rtsp://stream.isma.tv/foo.mp4 RTSP/1.0
CSeq: 6
Range: npt=3.676667-7.200000
Session: 2160059768067155308
User-Agent: ISMA-Client


RTSP/1.0 200 OK
Server: ISMA-Server
Cseq: 6
Session: 2160059768067155308
RTP-Info: url=trackID=6;seq=18562;rtptime=1426707168
 ,url=trackID=5;seq=27853;rtptime=1270047053


TEARDOWN rtsp://stream.isma.tv/foo.mp4 RTSP/1.0
CSeq: 7
Session: 2160059768067155308
User-Agent: ISMA-Client

RTSP/1.0 200 OK
Server: ISMA-Server
Cseq: 7
Session: 2160059768067155308
```

11 Appendix E: Minimal MPEG-4 Systems Support in ISMA (Normative)

11.1 Introduction

ISMA requires the support of MPEG-4 Systems. This occurs at two different places:

- in the MP4 file, in the IOD atom and BIFS and OD tracks, strictly following ISO/IEC 14496-12 and ISO/IEC 14496-14.
- in streaming, in the SDP.

11.2 MPEG-4 Systems content

11.2.1 The IOD

ISO/IEC 14496-1 (MPEG-4 Systems) requires the presence of an Initial Object Descriptor describing the BIFS and OD streams (when present). ISMA 1.0 content will therefore include the IOD.
Table E-1 shows the IOD of an ISMA 1.0 scene using the XMT format (XML-based format standardized by MPEG for authoring purposes).

Note: the values indicated in italic can be different as far as they respect the MPEG-4 systems specifications. Reasonable values are indicated in the table. Not all parameters are shown.

Table E-1 – IOD of an ISMA 1.0 scene

```
<InitialObjectDescriptor
        ODProfileLevelIndication= the proper profile indication
        SceneProfileLevelIndication= the proper profile indication
        AudioProfileLevelIndication= the proper profile indication
        VisualProfileLevelIndication= the proper profile indication
        graphicsProfileLevelIndication= the proper profile indication
>
        <esDescr>
                <ES_Descriptor ES_ID="1">
                        <decConfigDescr>
                                <DecoderConfigDescriptor
                                        objectTypeIndication="1"
                                        streamType="1"
                                >
                        </decConfigDescr>
                        <slConfigDescr>
                                <SLConfigDescriptor as appropriate />
                        </slConfigDescr>
                </ES_Descriptor>
                <ES_Descriptor ES_ID="2">
                        <decConfigDescr>
                                <DecoderConfigDescriptor
                                        objectTypeIndication="2"
                                        streamType="3"
                                >
                                        <decSpecificInfo>
                                                <BIFSv2Config as appropriate />
```

```
                                        </decSpecificInfo>
                                    /DecoderConfigDescriptor>
                                </decConfigDescr>
                                <slConfigDescr>
                                    <SLConfigDescriptor as appropriate />
                                </slConfigDescr>
                            </ES_Descriptor>
                        </esDescr>
</InitialObjectDescriptor>
```

11.2.2 The BIFS stream

ISMA 1.0 content is a presentation with a simple scene that includes at most one audio stream and one video stream. In BIFS terms, the scene contains a node for the audio object, and a rectangle node, whose texture is the video. In case the presentation includes only video or only audio, only one of these nodes would be present.
The following tables show (using VRML format for readability) the BIFS for the three possible scenes.

Note: the values indicated in italic can be different as far as they respect the MPEG-4 systems specifications. Reasonable values are indicated in the tables.

Table E-2 shows the description of an audio + video scene:

Table E-2 – BIFS representation of an ISMA 1.0 audio + video scene

```
OrderedGroup {
    children [
        Sound2D {
                spatialize FALSE
                source AudioSource {url 10}}
        Shape {
                geometry Bitmap { scale 1 1}
                appearance Appearance {texture MovieTexture {url 20}}
        }
    ]
}
```

Table E-3 shows the description of a video-only scene:

Table E-3 – BIFS representation of an ISMA 1.0 video-only scene

```
OrderedGroup {
    children [
        Shape {
                geometry Bitmap {scale 1 1}
                appearance Appearance {texture MovieTexture {url 20}}
        }
    ]
}
```

Table E-4 shows the description of an audio-only scene:

Table E-4 – BIFS representation of an ISMA 1.0 audio-only scene

```
OrderedGroup {
      children [
          Sound2D {
                spatialize FALSE
                source AudioSource {url 10}}
      ]
}
```

11.2.3 The OD stream

Given the simple BIFS scene of ISMA 1.0, the corresponding OD stream is made of just one *ObjectDescriptorUpdate* command that contains either one or two *ObjectDescriptors*; one for audio, one for video. Each of the ODs contains one ESD.

Table E-5 shows the OD stream of an ISMA 1.0 audio and video scene using the XMT format (XML-based format standardized by MPEG for authoring purposes).

Note: the values indicated in bold-italic can be different as far as they respect the MPEG-4 systems specifications. Reasonable values are indicated in the table. Not all parameters are shown.

Table E-5 – OD stream for ISMA 1.0 audio + video scene

```
< ObjectDescriptorUpdate>
        <OD>
                <ObjectDescriptor objectDescriptorID="10">
                        <esDescr>
                                <ES_Descriptor ES_ID="101">
                                        <decConfigDescr>
                                                <DecoderConfigDescriptor
                                                        objectTypeIndication="64"
                                                        streamType="5"
                                                >
                                                        <decSpecificInfo>
                                                                <AudioSpecificInfo audio
decoder config info goes here />
                                                        </decSpecificInfo>
                                                /DecoderConfigDescriptor>
                                        </decConfigDescr>
                                        <slConfigDescr>
                                                <SLConfigDescriptor as appropriate />
                                        </slConfigDescr>
                                </ES_Descriptor>
                        </esDescr>
                </ObjectDescriptor>
                <ObjectDescriptor objectDescriptorID="20">
                        <esDescr>
                                <ES_Descriptor ES_ID="201">
                                        <decConfigDescr>
                                                <DecoderConfigDescriptor
```

```
                                                objectTypeIndication="32"
                                                streamType="4"
                                            >
                                                <decSpecificInfo>
                                                    <VideoSpecificInfo video
decoder config info goes here />
                                                </decSpecificInfo>
                                            </DecoderConfigDescriptor>
                                        </decConfigDescr>
                                        <slConfigDescr>
                                            <SLConfigDescriptor as appropriate  />
                                        </slConfigDescr>
                                    </ES_Descriptor>
                                </esDescr>
                            </ObjectDescriptor>
                        </OD>
< /ObjectDescriptorUpdate>
```

Explanation: There is one *ObjectDescriptorUpdate* command that contains two *ObjectDescriptors*, one for audio and one for video, with ID *10* and *20* respectively. Each of the ODs contains one ESD, with ID *101* and *201* respectively.

The value of *objectTypeIndication* is, according to the standard, 64 for MPEG-4 audio and 32 for MPEG-4 video. Likewise *streamType* is 4 for video and 5 for audio.

The *decSpecificInfo* field within the OD structures might contain decoder specific info. See clause 11.3 for details.

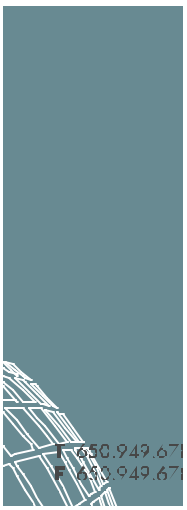
For the *SLConfigDescriptor* refer to clause 11.4.

11.3 Notes on the Decoder Specific Information

The *decSpecificInfo* field within the OD structures contains decoder specific info. There is Video specific info and audio specific info as defined in ISO/IEC 14496-2 and ISO/IEC 14496-3 respectively. E.g. for video, this contains the video headers – VOSH, VOL, etc. However in the streaming scenario the decSpecificInfo for video and audio streams need not be embedded in the OD stream, since it has to be in anycase exposed at the SDP level (in the config attribute of the format parameter). This implies that MPEG-4 Systems enabled receivers MUST be prepared to accept empty decSpecificInfo in the OD stream and retrieve the actual decSpecificInfo from the SDP. Moreover, in case the decSpecificInfo is also made available within the OD stream, it MUST coincide with that exposed at the SDP level.

11.4 Notes on the Sync Layer Config Descriptor

The *SLConfigDescriptor* contains parameters that are used to interpret the SL packet headers of the data. SL packets are described in ISO/IEC 14496-1, and provide a delivery-agnostic mechanism to associate meta-information such as timestamp, random access flags and so on, to the individual portions of the media payload. SL packets can be virtually obtained in both file-reading and streaming scenarios, by applying appropriate mapping rules from the corresponding meta-information within the MP4 file or in the RTP packet headers. This descriptor contains several fields whose value varies



between file-reading and streaming scenarios. Within MP4 files predefined=2 is used for all streams, implying certain values for the various individual fields, whereas in the streaming scenario ISMA has defined the individual field values which are appropriate for audio and video, and adopted predefined=1 for BIFS and OD – which represent a trivial case. For a detailed description of all fields refer to ISO/IEC 14496-1, clause 10.2.3 and 10.2.4.

11.5 Notes on Synchronization

11.5.1 The MPEG-4 Systems model for synchronization

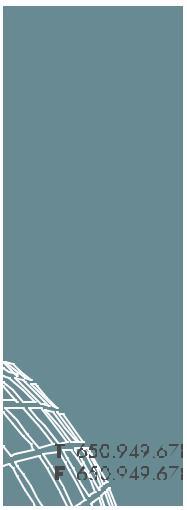
MPEG-4 Systems provides a sophisticated synchronization model, in which multiple timelines can be run concurrently each synchronizing a subset of elementary streams within the presentation. This model requires that each elementary stream is explicitly assigned to a particular timeline, and the OD stream, by means of the ES_Id, OCRstreamFlag, OCR_ES_Id fields in the ESDescriptor and the OCRlength and OCRresolution fields in the SLconfigDescriptor, provide the necessary information.
The OCRstreamFlag indicates whether an elementary stream uses the clock of another stream (indicated in the OCR_ES_Id field) or its own. The clock itself is obtained by conveying samples of the OCR (Object Clock Reference), whose resolution and length are provided in the OCRresolution and OCRlength fields of the SlconfigDescriptor of the elementary stream to which the clock is associated. In case OCRlength is 0, no actual samples of the OCR are carried, and the timeline is locally generated: this is the case for the file-reading scenario. Still, even in this case, the OCRstreamFlag and OCR_ES_Id fields allow syncing elementary stream, as well as running multiple timelines.
In the case of ISMA1.0 content this model is overkill, since all streams shall be tight to a single timeline. This is true for both the file-reading and the streaming scenarios.

11.5.2 File-reading scenario

In the file-reading scenario, all tracks in the MP4 file must be synchronized (including BIFS & OD tracks). The simplest way to indicate this is to not place any 'sync' track reference atom in the file (this means that all tracks are synchronized by default). Also, the SlconfigDescriptor will have predefined=2, which implies OCRlength=0.

11.5.3 Streaming scenario

In the streaming scenario the RTP mechanism for synchronization is used. Players that depend on the MPEG-4 Systems model for synchronization, shall be able to map the RTP mechanism to the OCR model. Specifically, the server time reported through Sender Report commands in RTCP will be translated into OCRs when the player performs SL mapping. In presentations that include video and audio, it is recommended to assume that the OCRs are conveyed with one stream, preferably audio, with the other stream depending on it synchronization. Table E-7 follows this model, and assumes that the audio stream is the one that carries the OCRs, while the video stream clock depends on the audio stream. If this pattern does not apply, for instance, if the video stream plays continuously and the audio stream is sometime muted, the values in the table should be changed accordingly. Specifically, OCRstreamFlag should be set to 1 if the stream clock depends on another stream, and, in this case, and only in this case, the ES_Descriptor of the stream shall include an OCR_ES_Id field that contains the ES_id of the stream that carries the OCRs.



11.6 Notes on Timescales

OD and SDP both provide the description of the timescale by which the values of timestamps (and clock references) are measured. This might generate some confusion. There is no reason for setting the two timescales to different values: a discrepancy would not prevent interoperability, however it is recommended to set both timescales identically. In any case the bits-on-the-wire only reflect the timescale value as indicated in SDP.

11.7 MPEG-4 Systems support in MP4 files

ISMA 1.0 strictly adheres to the ISO/IEC 14496-12 (ISO Base Media File Format) and ISO/IEC 14496-14 (MP4 File Format).
In particular the “iod ” box, the BIFS track, the OD track, the “esds” boxes for the Audio and Video tracks, are fully conformant to those specifications.
Note that the SLConfigDescriptor for all streams are set to “predefined=2”.

11.8 MPEG-4 Systems support in SDP

11.8.1 The self-contained IOD

MPEG-4 Systems requires that the scene description and the object description are conveyed in separate BIFS and OD streams. However, in order to avoid the consumption of additional resources in the streaming scenario, ISMA 1.0 requires that these streams be embedded within the IOD that is then conveyed as part of the SDP description.

This mechanism is based on the following points:

- MPEG-4 allows using URLs in ESDs
- A URL can be of type “data:”, i.e. the content pointed to by the URL can be embedded in the URL itself.

Using these principles, Table E-6 shows the self-contained IOD of an ISMA 1.0 presentation using the XMT format: the only difference with respect to the IOD shown in Table E-1 is given by the URLString fields containing the “data:” URLs.

Note: the values indicated in bold-italic can be different as far as they respect the MPEG-4 systems specifications. Reasonable values are indicated in the table. Not all parameters are shown.

Table E-6 – A self-contained IOD for the simple scene

```
<InitialObjectDescriptor
        ODProfileLevelIndication= the proper profile indication
        SceneProfileLevelIndication= the proper profile indication
        AudioProfileLevelIndication= the proper profile indication
        VisualProfileLevelIndication= the proper profile indication
        graphicsProfileLevelIndication= the proper profile indication
>
        <esDescr>
                <ES_Descriptor ES_ID="1"
                        URLString=”data:application/mpeg4-od-au;base64,<Base64 encoded OD
stream >”
                >
                        <decConfigDescr>
```

```
                        <DecoderConfigDescriptor
                                objectTypeIndication="1"
                                streamType="1"
                        >
                    </decConfigDescr>
                    <slConfigDescr>
                        <SLConfigDescriptor predefined=”1” />
                    </slConfigDescr>
                </ES_Descriptor>
                <ES_Descriptor ES_ID="2"
                    URLString=”data:application/mpeg4-bifs-au;base64,<Base64 encoded BIFS
stream >”
                >
                    <decConfigDescr>
                        <DecoderConfigDescriptor
                                objectTypeIndication="2"
                                streamType="3"
                        >
                            <decSpecificInfo>
                                <BIFSv2Config as appropriate />
                            </decSpecificInfo>
                        /DecoderConfigDescriptor>
                    </decConfigDescr>
                    <slConfigDescr>
                        <SLConfigDescriptor predefined=”1” />
                    </slConfigDescr>
                </ES_Descriptor>
        </esDescr>
</InitialObjectDescriptor>
```

The code in Table E-6 assumes that “data:” URLs contain one access unit with null SL packet header.

11.8.2 MPEG-4 Systems Information in SDP

The IOD shall be included in the SDP description, using the *mpeg4-iod* parameter described in ISO/IEC 14496-8 as:

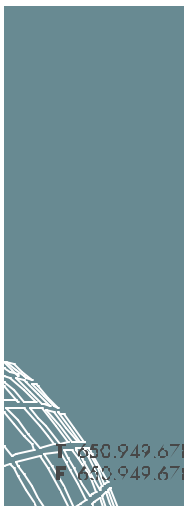
a=mpeg4-iod [<location>*]*

The *location* shall be a URL enclosed in double-quotes, which will supply the IOD in its standard binary format. The IOD may be embedded in a "data:" URL and Base64 encoding of binary data (described in RFC 1341).

The media streams shall be described in the SDP using the usual parameters. In addition, there is a need to associate ESIDs with the corresponding stream description. This is done as described in ISO/IEC 14496-8, i.e. a stream-specific attribute shall be present for each MPEG-4 stream. The attribute will take the following form:

a=mpeg4-esid esid

where *esid* is the ESID.



12 Appendix F: Example of Minimal MPEG-4 Systems Support in ISMA (Informative)

The examples below are based on the Tables of Appendix E, with specific values assigned to variable fields. The examples specifically address the streaming scenario and show:

- the binary representation of BIFS bitstreams for Audio+Video, VideoOnly and AudioOnly scenes;
- the binary representation of the corresponding OD bitstreams, for both ISMA profile0 and profile1;
- the binary representation of IODs, embedding BIFS and OD bitstreams by means of “data:” URLs, for both ISMA profile0 and profile1.

These are examples and by no means restrict the set of valid BIFS/OD/IOD that can be used. An MPEG-4 Systems unaware content producer / streamer can however safely pick from these examples in order to generate the IOD corresponding to the appropriate ISMA Profile and scene type.

Note that the DecoderSpecificInfo descriptors of the audio and video streams are omitted, following the rule specified in Appendix E that DecoderSpecificInfo may be omitted if provided in other means.
Note also that the tables below refer to the OD and IOD used in the streaming context, not to those stored in the file and used for local playback. More specifically the SLConfigDescriptor when stored in the file contains the value predefined=2 and no other fields. Also, several other IOD/OD structures are mapped into specific structures of the MP4 file, e.g. the Elementary Stream Descriptor is stored as part of the Sample Description Box.

12.1 BIFS bitstreams

12.1.1 ISMA1.0 audio+video BIFS

The binary representation of the scene in Table E-2 is the following:

```
TOTAL OF 24 BYTES
C0 10 12 81 93 02 A0 57 26 10 41 FC 00 00 01 FC 00 00 04 42 82 28
29 F8
```

Which in base64 becomes:

```
wBASgZMCoFcmEEH8AAAB/AAABEKCKCn4
```

12.1.2 ISMA1.0 video-only BIFS

The binary representation of the scene in Table E-3 is the following:

```
TOTAL OF 19 BYTES
C0 10 12 61 04 1F C0 00 00 1F C0 00 00 44 28 22 82 9F 80
```

Which in base64 becomes:

```
wBASYQQfwAAAH8AAAEQoIoKfgA==
```

12.1.3 ISMA1.0 audio-only BIFS

The binary representation of the scene in Table E-4 is the following:

```
TOTAL OF 9 BYTES
C0 10 12 81 93 02 A0 57 C0
```

Which in base64 becomes:

```
wBASgZMCoFfA
```

12.2 OD bitstreams

12.2.1 The complete representation of the OD

The OD bitstreams for the various combinations of ISMA profiles and scene types differ for just a very few elements:

- the presence of the Audio and/or Video Object Descriptor
- the fields related to the clock reference stream (in case of Audio+Video scene, the Video depends on the Audio, otherwise no dependency exists)
- the values in the decoder config descriptor (wrt: profile 0 and profile 1)
- the values in the SL config descriptor (timestamp/OCRResolution, timestamp/OCRLength)
- the overall descriptor sizes

Table E-7 provides in its 6 parallel columns the actual values for the 6 possible combinations, namely Audio+Video, Video-only and Audio-only scenes for both ISMA Profiles 0 and 1.

This table specifically represents the OD bitstream for the streaming scenario, however, given the similarity with the OD representation within the MP4 file, the few field values (but for the descriptor size values) that would vary in this second case are indicated in parenthesis (e.g. field 5.3). Therefore, if the field contains a value in parenthesis '(' and ')', the value in parenthesis is the value for the file representation: if the value in parenthesis is a '-', it is not present in the file representation.

Note: the values indicated in italic between ‘<’ and ‘>’ can be different as far as they respect the MPEG-4 systems specifications (e.g. field 2.3). Reasonable values are indicated in the table. Some descriptor sizes too are indicated between ‘<’ and ‘>’, since their values may be affected by the presence/absence of certain fields, specifically the OCR_ES_Id fields (3.8 and 7.8: either one or the other must be present for the AV scene in the streaming scenario).

Table E-7 – Full representation of OD bitstreams for ISMA 1.0 presentations

Field No.	Size in Bits	Field Name	Values for Audio and/or Video scenes, profiles 0 or 1					
			AV 0	V 0	A 0	AV 1	V 1	A 1
ObjectDescriptor Access Unit								
1.1	8	ObjectDescriptorUpdate tag	1					
1.2	8	Descriptor size	86	42	42	86	42	42
ObjectDescriptor (of Video stream)								

2.1	8	ObjectDescriptor tag	1						
2.2	8	Descriptor size	*<42>*	40	-	*<42>*	40	-	
2.3	10	ObjectDescriptorID	*<20>*						
2.4	1	URL_Flag	0						
2.5	5	Reserved	31						
ES_Descriptor (of Video stream)									
3.1	8	ES_Descriptor tag	3						
3.2	8	Descriptor size	*<38>*	*36*	-	*<38>*	36	-	
3.3	16	ES_ID	*<201>*						
3.4	1	streamDependenceFlag	0						
3.5	1	URL_Flag	0						
3.6	1	OCRstreamFlag	*<1>*	0	-	*<1>*	0	-	
3.7	5	streamPriority	*<0>*						
3.8	16	OCR_ES_Id	*<101>*	-	-	*<101>*	-	-	
DecoderConfigDescriptor (of Video stream)									
4.1	8	DecoderConfigDescriptor tag	4						
4.2	8	Descriptor size	13						
4.3	8	objectTypeIndication	32						
4.4	6	streamType	4						
4.5	1	upStream	0						
4.6	1	Reserved	1						
4.7	24	bufferSizeDB [1]	*<20480>*	*<20480>*	-	*<133120>*	*<133120>*	-	
4.8	32	maxBitrate [1]	*<64000>*	*<64000>*	-	*<1500000>*	*<1500000>*	-	
4.9	32	avgBitrate [1]	*<64000>*	*<64000>*	-	*<1500000>*	*<1500000>*	-	
SLConfigDescriptor (of Video stream)									
5.1	8	SLConfigDescriptor tag	6						
5.2	8	Descriptor size	16						
5.3	8	predefined	0 (2)						
5.4	1	useAccessUnitStartFlag	0 (-)						
5.5	1	useAccessUnitEndFlag	1 (-)						
5.6	1	useRandomAccessPointFlag	0 (-)						
5.7	1	hasRandomAccessUnitsOnlyFlag	0 (-)						
5.8	1	usePaddingFlag	0 (-)						
5.9	1	useTimeStampsFlag	1 (-)						
5.10	1	useIdleFlag	0 (-)						
5.11	1	durationFlag	0 (-)						
5.12	32	timeStampResolution	*see note [2]*						
5.13	32	OCRResolution	*<0>(-)*						
5.14	8	timeStampLength	*<32>* (-)						
5.15	8	OCRLength	*<0>* (-)						
5.16	8	AU_Length	0 (-)						
5.17	8	instantBitrateLength	0 (-)						

[1] The values indicated as default are the maximum values allowed for the profile. Note that implementations might benefit from more accurate computations.

[2] For the streaming scenario, it is recommended to use the same value as provided in the corresponding rtpmap line of the SDP description. For the file-reading scenario, this is the timescale value as indicated in the corresponding 'mdhd' (Media Header) box. Note that there is no reason for having discrepancies among these values.

5.18	4	degradationPriorityLength	0 (-)						
5.19	5	AU_seqNumLength	0 (-)						
5.20	5	packetSeqNumLength	0 (-)						
5.21	2	reserved	3						
ObjectDescriptor (of Audio stream)									
6.1	8	ObjectDescriptor tag	1						
6.2	8	Descriptor size	*<40>*	-	40	*<40>*	-	40	
6.3	10	ObjectDescriptorID	*<10>*						
6.4	1	URL_Flag	0						
6.5	5	Reserved	31						
ES_Descriptor (of Audio stream)									
7.1	8	ES_Descriptor tag	3						
7.2	8	Descriptor size	<36>	-	36	<36>	-	36	
7.3	16	ES_ID	*<101>*						
7.4	1	streamDependenceFlag	0						
7.5	1	URL_Flag	0						
7.6	1	OCRstreamFlag	*<0>*	0	-	*<0>*	0	-	
7.7	5	streamPriority	*<0>*						
7.8	16	OCR_ES_Id	<->	-	-	<->	-	-	
DecoderConfigDescriptor (of Audio stream)									
8.1	8	DecoderConfigDescriptor tag	4						
8.2	8	Descriptor size	13						
8.3	8	objectTypeIndication	64						
8.4	6	streamType	5						
8.5	1	upStream	0						
8.6	1	Reserved	1						
8.7	24	bufferSizeDB [1]	*<8000>*	-	*<8000>*	*<8000>*	-	*<8000>*	
8.8	32	maxBitrate [1]	*<128000>*	-	*<128000>*	*<128000>*	-	*<128000>*	
8.9	32	avgBitrate [1]	*<128000>*	-	*<128000>*	*<128000>*	-	*<128000>*	
SLConfigDescriptor (of Audio stream)									
9.1	8	SLConfigDescriptor tag	6						
9.2	8	Descriptor size	16						
9.3	8	predefined	0 (2)						
9.4	1	useAccessUnitStartFlag	0 (-)						
9.5	1	useAccessUnitEndFlag	1 (-)						
9.6	1	useRandomAccessPointFlag	0 (-)						
9.7	1	hasRandomAccessUnitsOnlyFlag	0 (-)						
9.8	1	usePaddingFlag	0 (-)						
9.9	1	useTimeStampsFlag	1 (-)						
9.10	1	useIdleFlag	0 (-)						
9.11	1	durationFlag	0 (-)						
9.12	32	timeStampResolution	*see note [2]*						
9.13	32	OCRResolution	*<1000>(-)*						
9.14	8	timeStampLength	*<32>* (-)						
9.15	8	OCRLength	*<32>* (-)						
9.16	8	AU_Length	0 (-)						
9.17	8	instantBitrateLength	0 (-)						
9.18	4	degradationPriorityLength	0 (-)						

9.19	5	AU_seqNumLength	0 (-)
9.20	5	packetSeqNumLength	0 (-)
9.21	2	reserved	3

12.2.2 ISMA1.0 audio+video OD, profile 0
The binary representation of this bitstream is the following (assuming timeStampResolution=1000 and timeStampLength=32 for the audio and video ESDs):

```
TOTAL OF 88 BYTES
01 56 01 2a 05 1f 03 26 00 c9 20 00 65 04 0d 20 11 00 50 00 00 00
fa 00 00 00 fa 00 06 10 00 44 00 00 03 e8 00 00 00 00 20 00 00 00
00 03 01 28 02 9f 03 24 00 65 00 04 0d 40 15 00 1f 40 00 01 f4 00
00 01 f4 00 06 10 00 44 00 00 03 e8 00 00 03 e8 20 20 00 00 00 03
```

Which in base64 becomes:

```
AVYBKgUfAyYAySAAZQQNIBEAUAAAAPoAAAD6AAYQAEQAAAPoAAAAACAAAAAAAwEoA
p8DJABlAAQNQBUAH0AAAfQAAAH0AAYQAEQAAAPoAAAD6CAgAAAAAw==
```

12.2.3 ISMA1.0 video-only OD, profile 0
The binary representation of this bitstream is the following (assuming timeStampResolution=1000 and timeStampLength=32 for the video ESD):

```
TOTAL OF 44 BYTES
01 2a 01 28 05 1f 03 24 00 c9 00 04 0d 20 11 00 50 00 00 00 fa 00
00 00 fa 00 06 10 00 44 00 00 03 e8 00 00 00 00 20 00 00 00 00 03
```

Which in base64 becomes:

```
ASoBKAUfAyQAyQAEDSARAFAAAAD6AAAA+gAGEABEAAAD6AAAAAAgAAAAAAM=
```

12.2.4 ISMA1.0 audio-only OD, profile 0
The binary representation of this bitstream is the following (assuming timeStampResolution=1000 and timeStampLength=32 for the audio ESD):

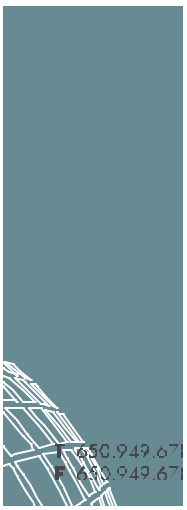
```
TOTAL OF 44 BYTES
01 2a 01 28 02 9f 03 24 00 65 00 04 0d 40 15 00 1f 40 00 01 f4 00
00 01 f4 00 06 10 00 44 00 00 03 e8 00 00 03 e8 20 20 00 00 00 03
```

Which in base64 becomes:

```
ASoBKAKfAyQAZQAEDUAVAB9AAAH0AAAB9AAGEABEAAAD6AAAA+ggIAAAAAM=
```

12.2.5 ISMA1.0 audio+video OD, profile 1
The binary representation of this bitstream is the following (assuming timeStampResolution=1000 and timeStampLength=32 for the audio and video ESDs):

```
TOTAL OF 88 BYTES
01 56 01 2a 05 1f 03 26 00 c9 20 00 65 04 0d 20 11 02 08 00 00 16
e3 60 00 16 e3 60 06 10 00 44 00 00 03 e8 00 00 00 00 20 00 00 00
```

```
00 03 01 28 02 9f 03 24 00 65 00 04 0d 40 15 00 1f 40 00 01 f4 00
00 01 f4 00 06 10 00 44 00 00 03 e8 00 00 03 e8 20 20 00 00 00 03
```

Which in base64 becomes:

```
AVYBKgUfAyYAySAAZQQNIBECCAAAFuNgABbjYAYQAEQAAAPoAAAAACAAAAAAAwEoA
p8DJABlAAQNQBUAH0AAAfQAAAH0AAYQAEQAAAPoAAAD6CAgAAAAAw==
```

12.2.6 ISMA1.0 video-only OD, profile 1
The binary representation of this bitstream is the following (assuming timeStampResolution=1000 and timeStampLength=32 for the video ESD):

```
TOTAL OF 44 BYTES
01 2a 01 28 05 1f 03 24 00 c9 00 04 0d 20 11 02 08 00 00 16 e3 60
00 16 e3 60 06 10 00 44 00 00 03 e8 00 00 00 00 20 00 00 00 00 03
```

Which in base64 becomes:

```
ASoBKAUfAyQAyQAEDSARAggAABbjYAAW42AGEABEAAAD6AAAAAAgAAAAAAM=
```

12.2.7 ISMA1.0 audio-only OD, profile 1
The binary representation of this bitstream is identical to that of profile 0

12.3 IODs

12.3.1 The complete representation of the IOD
The IODs for the various combinations of ISMA profiles and scene types differ for just a very few elements:

- the profileLevelIndications
- the embedded BIFS and OD “data:” URLs
- the overall descriptor sizes

Table E-8 provides in its 6 parallel columns the actual values for the 6 possible combinations, namely Audio+Video, Video-only and Audio-only scenes for both ISMA Profiles 0 and 1.

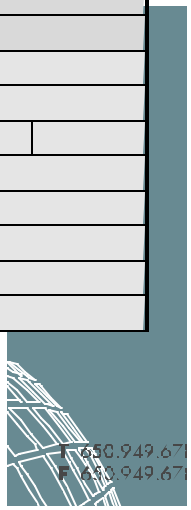
This table specifically represents the IOD for the streaming scenario, however, given the similarity with the IOD representation within the MP4 file, the few field values (but for the descriptor size values) that would vary in this second case are indicated in parenthesis (e.g. field 2.5). Therefore, if the field contains a value in parenthesis '(' and ')', the value in parenthesis is the value for the file representation: if the value in parenthesis is a '-', it is not present in the file representation.

Note: the values indicated in italic between ‘<’ and ‘>’ can be different as far as they respect the MPEG-4 systems specifications (e.g. field 1.3). Reasonable values are indicated in the table.

Table E-8 – Full representation of IODs for ISMA 1.0 presentations

Field	Size in	Field Name	Values for Audio and/or Video scenes, profiles 0 or 1

No.	Bits		AV 0	V 0	A 0	AV 1	V 1	A 1
		InitialObjectDescriptor						
1.1	8	InitialObjectDescriptor tag	2					
1.2	16	Descriptor size						
1.3	10	ObjectDescriptorID	*<1>*					
1.4	1	URL_Flag	0					
1.5	1	includeInlineProfilesFlag	0					
1.6	4	Reserved	15					
1.7	8	ODProfileLevelIndication	255					
1.8	8	sceneProfileLevelIndication	255					
1.9	8	audioProfileLevelIndication	*<15>*	255	*<15>*	*<15>*	255	*<15>*
1.10	8	visualProfileLevelIndication	*<1>*	*<1>*	255	*<247>*	*<247>*	255
1.11	8	graphicsProfileLevelIndication	255					
		ES_Descriptor (of OD stream)						
2.1	8	ES_Descriptor tag	3					
2.2	8	Descriptor size						
2.3	16	ES_ID	*<1>*					
2.4	1	streamDependenceFlag	0					
2.5	1	URL_Flag	1 (0)					
2.6	1	OCRstreamFlag	0					
2.7	5	streamPriority	*<0>*					
2.8	8	string-size						
2.9	36*8	URLstring (common part)	"data:application/mpeg4-od-au;base64," (-)					
2.10	n*8	URLstring (specific part)	Based on the specific OD AU (-)					
		DecoderConfigDescriptor (of OD stream)						
3.1	8	DecoderConfigDescriptor tag	4					
3.2	8	Descriptor size	13					
3.3	8	objectTypeIndication	1					
3.4	6	streamType	1					
3.5	1	upStream	0					
3.6	1	Reserved	1					
3.7	24	bufferSizeDB	*<200>*					
3.8	32	maxBitrate	*<0>*					
3.9	32	avgBitrate	*<0>*					
		SLConfigDescriptor (of OD stream)						
4.1	8	SLConfigDescriptor tag	6					
4.2	8	Descriptor size	9					
4.3	8	predefined	1 (2)					
4.4	32	startDecodingTimeStamp	0 (-)					
4.5	32	startCompositionTimeStamp	0 (-)					
		ES_Descriptor (of BIFS stream)						
5.1	8	ES_Descriptor tag	3					
5.2	8	Descriptor size						
5.3	16	ES_ID	*<2>*					
5.4	1	StreamDependenceFlag	0					
5.5	1	URL_Flag	1 (0)					
5.6	1	OCRstreamFlag	0					
5.7	5	StreamPriority	*<0>*					



5.8	8	string-size							
5.9	38*8	URLstring (common part)	"data:application/mpeg4-bifs-au;base64," (-)						
5.10	n*8	URLstring (specific part)	Based on the specific BIFS AU (-)						
DecoderConfigDescriptor (of BIFS)									
6.1	8	DecoderConfigDescriptor tag	4						
6.2	8	Descriptor size	18						
6.3	8	ObjectTypeIndication	2						
6.4	6	StreamType	3						
6.5	1	Upstream	0						
6.6	1	Reserved	1						
6.7	24	BufferSizeDB	*<100>*						
6.8	32	MaxBitrate	*<0>*						
6.9	32	AvgBitrate	*<0>*						
BIFSv2Config									
7.1	8	BIFSv2Config tag	5						
7.2	8	Descriptor size	3						
7.3	1	use3DmeshCoding	0						
7.4	1	UsePredictiveMFField	0						
7.5	5	NodeIDbits	0						
7.6	5	RouteIDbits	0						
7.7	5	ProtoIDbits	0						
7.8	1	IsCommandStream	1						
7.9	1	PixelMetric	1						
7.10	1	HasSize	0						
7.11	4	byte align	0						
SLConfigDescriptor (of BIFS)									
8.1	8	SLConfigDescriptor tag	6						
8.2	8	descriptor size	9						
8.3	8	Predefined	1 (2)						
8.4	32	startDecodingTimeStamp	0 (-)						
8.5	32	startCompositionTimeStamp	0 (-)						

12.3.2 ISMA1.0 audio+video IOD, profile 0

The binary representation of the IOD comprising the BIFS and OD bitstreams as provided in 12.1.1 and 12.2.2 is the following:

```
TOTAL OF 306 BYTES
02 82 2f 00 4f ff ff 0f 01 ff 03 81 3a 00 01 40 9c 64 61 74 61 3a
61 70 70 6c 69 63 61 74 69 6f 6e 2f 6d 70 65 67 34 2d 6f 64 2d 61
75 3b 62 61 73 65 36 34 2c 41 56 59 42 4b 67 55 66 41 79 59 41 79
53 41 41 5a 51 51 4e 49 42 45 41 55 41 41 41 41 50 6f 41 41 41 44
36 41 41 59 51 41 45 51 41 41 41 50 6f 41 41 41 41 41 43 41 41 41
41 41 41 41 77 45 6f 41 70 38 44 4a 41 42 6c 41 41 51 4e 51 42 55
41 48 30 41 41 41 66 51 41 41 41 48 30 41 41 59 51 41 45 51 41 41
41 50 6f 41 41 41 44 36 43 41 67 41 41 41 41 41 77 3d 3d 04 0d 01
05 00 00 c8 00 00 00 00 00 00 00 00 06 09 01 00 00 00 00 00 00 00
00 03 69 00 02 40 46 64 61 74 61 3a 61 70 70 6c 69 63 61 74 69 6f
6e 2f 6d 70 65 67 34 2d 62 69 66 73 2d 61 75 3b 62 61 73 65 36 34
2c 77 42 41 53 67 5a 4d 43 6f 46 63 6d 45 45 48 38 41 41 41 42 2f
```

```
41 41 41 42 45 4b 43 4b 43 6e 34 04 12 02 0d 00 00 64 00 00 00 00
00 00 00 00 05 03 00 00 60 06 09 01 00 00 00 00 00 00 00 00
```

Which in base64 becomes:

```
AoIvAE///w8B/wOBOgABQJxkYXRhOmFwcGxpY2F0aW9uL21wZWc0LW9kLWF1O2Jhc
2U2NCxBVllCS2dVZkF5WUF5U0FBWlFRTklCRUFVQUFBQVBvQUFBRDZBQVlRQUVRQU
FBUG9BQUFBQUNBQUFBQUFBd0VvQXA4REpBQmxBQVFOUUJVQUgwQUFBZlFBQUFIMEF
BWVFBRVFBQUFQb0FBQUQ2Q0FnQUFBQUF3PT0EDQEFAADIAAAAAAAAAAAGCQEAAAAA
AAAAAANpAAJARmRhdGE6YXBwbGljYXRpb24vbXBlZzQtYmlmcy1hdTtiYXNlNjQsd
0JBU2daTUNvRmNtRUVIOEFBQUIvQUFBQkVLQ0tDbjQEEgINAABkAAAAAAAAAAAFAw
AAYAYJAQAAAAAAAAA
```

12.3.3 ISMA1.0 video-only IOD, profile 0

The binary representation of the IOD comprising the BIFS and OD bitstreams as provided in 12.1.2 and 12.2.3 is the following:

```
TOTAL OF 241 BYTES
02 81 6e 00 4f ff ff ff 01 ff 03 7e 00 01 40 60 64 61 74 61 3a 61
70 70 6c 69 63 61 74 69 6f 6e 2f 6d 70 65 67 34 2d 6f 64 2d 61 75
3b 62 61 73 65 36 34 2c 41 53 6f 42 4b 41 55 66 41 79 51 41 79 51
41 45 44 53 41 52 41 46 41 41 41 41 44 36 41 41 41 41 2b 67 41 47
45 41 42 45 41 41 41 44 36 41 41 41 41 41 41 67 41 41 41 41 41 41
4d 3d 04 0d 01 05 00 00 c8 00 00 00 00 00 00 00 00 06 09 01 00 00
00 00 00 00 00 00 03 65 00 02 40 42 64 61 74 61 3a 61 70 70 6c 69
63 61 74 69 6f 6e 2f 6d 70 65 67 34 2d 62 69 66 73 2d 61 75 3b 62
61 73 65 36 34 2c 77 42 41 53 59 51 51 66 77 41 41 41 48 38 41 41
41 45 51 6f 49 6f 4b 66 67 41 3d 3d 04 12 02 0d 00 00 64 00 00 00
00 00 00 00 00 05 03 00 00 60 06 09 01 00 00 00 00 00 00 00 00
```

Which in base64 becomes:

```
AoFuAE////8B/wN+AAFAYGRhdGE6YXBwbGljYXRpb24vbXBlZzQtb2QtYXU7YmFzZ
TY0LEFTb0JLQVVmQXlRQXlRQUVEU0FSQUZBQUFBRDZBQUFBK2dBR0VBQkVBQUFENk
FBQUFBQWdBQUFBQUFNPQQNAQUAAMgAAAAAAAAAAAYJAQAAAAAAAAAA2UAAkBCZGF
0YTphcHBsaWNhdGlvbi9tcGVnNC1iaWZzLWF1O2Jhc2U2NCx3QkFTWVFRZndBQUFI
OEFBQUVRb0lvS2ZnQT09BBICDQAAZAAAAAAAAAAABQMAAGAGCQEAAAAAAAAAA==
```

12.3.4 ISMA1.0 audio-only IOD, profile 0

The binary representation of the IOD comprising the BIFS and OD bitstreams as provided in 12.1.3 and 12.2.4 is the following:

```
TOTAL OF 225 BYTES
02 81 5e 00 4f ff ff 0f ff ff 03 7e 00 01 40 60 64 61 74 61 3a 61
70 70 6c 69 63 61 74 69 6f 6e 2f 6d 70 65 67 34 2d 6f 64 2d 61 75
3b 62 61 73 65 36 34 2c 41 53 6f 42 4b 41 4b 66 41 79 51 41 5a 51
41 45 44 55 41 56 41 42 39 41 41 41 48 30 41 41 41 42 39 41 41 47
45 41 42 45 41 41 41 44 36 41 41 41 41 2b 67 67 49 41 41 41 41 41
4d 3d 04 0d 01 05 00 00 c8 00 00 00 00 00 00 00 00 06 09 01 00 00
00 00 00 00 00 00 03 55 00 02 40 32 64 61 74 61 3a 61 70 70 6c 69
63 61 74 69 6f 6e 2f 6d 70 65 67 34 2d 62 69 66 73 2d 61 75 3b 62
```

```
61 73 65 36 34 2c 77 42 41 53 67 5a 4d 43 6f 46 66 41 04 12 02 0d
00 00 64 00 00 00 00 00 00 00 00 05 03 00 00 60 06 09 01 00 00 00
00 00 00 00 00
```

Which in base64 becomes:

```
AoFeAE///w///wN+AAFAYGRhdGE6YXBwbGljYXRpb24vbXBlZzQtb2QtYXU7YmFzZ
TY0LEFTb0JLQUtmQXlRQVpRQUVEVUFWQUI5QUFBSDBBQUFCOUFBR0VBQkVBQUFENk
FBQUErZ2dJQUFBQUFNPQQNAQUAAMgAAAAAAAAAAYJAQAAAAAAAAAA1UAAkAyZGF
0YTphcHBsaWNhdGlvbi9tcGVnNC1iaWZzLWF1O2Jhc2U2NCx3QkFTZ1pNQ29GZkEE
EgINAABkAAAAAAAAAAFAwAAYAYJAQAAAAAAAAA
```

12.3.5 ISMA1.0 audio+video IOD, profile 1

The binary representation of the IOD comprising the BIFS and OD bitstreams as provided in 12.1.1 and 12.2.2 is the following:

```
TOTAL OF 306 BYTES
02 82 2f 00 4f ff ff 0f f7 ff 03 81 3a 00 01 40 9c 64 61 74 61 3a
61 70 70 6c 69 63 61 74 69 6f 6e 2f 6d 70 65 67 34 2d 6f 64 2d 61
75 3b 62 61 73 65 36 34 2c 41 56 59 42 4b 67 55 66 41 79 59 41 79
53 41 41 5a 51 51 4e 49 42 45 43 43 41 41 41 46 75 4e 67 41 42 62
6a 59 41 59 51 41 45 51 41 41 41 50 6f 41 41 41 41 41 43 41 41 41
41 41 41 41 77 45 6f 41 70 38 44 4a 41 42 6c 41 41 51 4e 51 42 55
41 48 30 41 41 41 66 51 41 41 41 48 30 41 41 59 51 41 45 51 41 41
41 50 6f 41 41 41 44 36 43 41 67 41 41 41 41 41 77 3d 3d 04 0d 01
05 00 00 c8 00 00 00 00 00 00 00 00 06 09 01 00 00 00 00 00 00 00
00 03 69 00 02 40 46 64 61 74 61 3a 61 70 70 6c 69 63 61 74 69 6f
6e 2f 6d 70 65 67 34 2d 62 69 66 73 2d 61 75 3b 62 61 73 65 36 34
2c 77 42 41 53 67 5a 4d 43 6f 46 63 6d 45 45 48 38 41 41 41 42 2f
41 41 41 42 45 4b 43 4b 43 6e 34 04 12 02 0d 00 00 64 00 00 00 00
00 00 00 00 05 03 00 00 60 06 09 01 00 00 00 00 00 00 00 00 00
```

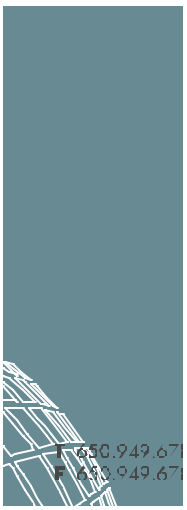
Which in base64 becomes:

```
AoIvAE///w/3/wOBOgABQJxkYXRhOmFwcGxpY2F0aW9uL21wZWc0LW9kLWF1O2Jhc
2U2NCxBVllCS2dVZkF5WUF5U0FBWlFRTklCRUNDQUFBRnVOZ0FCYmpZQVlRQUVRQU
FBUG9BQUFBQUNBQUFBQUFBd0VvQXA4REpBQmxBQVFOUUJVQUgwQUFBZlFBQUFIMEF
BWVFBRVFBQUFQb0FBQUQ2Q0FnQUFBQUF3PT0EDQEFAADIAAAAAAAAAAAGCQEAAAAA
AAAAAANpAAJARmRhdGE6YXBwbGljYXRpb24vbXBlZzQtYmlmcy1hdTtiYXNlNjQsd
0JBU2daTUNvRmNtRUVIOEFBQUIvQUFBQkVLQ0tDbjQEEgINAABkAAAAAAAAAAFAw
AAYAYJAQAAAAAAAAA
```

12.3.6 ISMA1.0 video-only IOD, profile 1

The binary representation of the IOD comprising the BIFS and OD bitstreams as provided in 12.1.2 and 12.2.3 is the following:

```
TOTAL OF 241 BYTES
02 81 6e 00 4f ff ff ff f7 ff 03 7e 00 01 40 60 64 61 74 61 3a 61
70 70 6c 69 63 61 74 69 6f 6e 2f 6d 70 65 67 34 2d 6f 64 2d 61 75
3b 62 61 73 65 36 34 2c 41 53 6f 42 4b 41 55 66 41 79 51 41 79 51
41 45 44 53 41 52 41 67 67 41 41 42 62 6a 59 41 41 57 34 32 41 47
```

```
45 41 42 45 41 41 41 44 36 41 41 41 41 41 41 67 41 41 41 41 41 41
4d 3d 04 0d 01 05 00 00 c8 00 00 00 00 00 00 00 00 06 09 01 00 00
00 00 00 00 00 00 03 65 00 02 40 42 64 61 74 61 3a 61 70 70 6c 69
63 61 74 69 6f 6e 2f 6d 70 65 67 34 2d 62 69 66 73 2d 61 75 3b 62
61 73 65 36 34 2c 77 42 41 53 59 51 51 66 77 41 41 41 48 38 41 41
41 45 51 6f 49 6f 4b 66 67 41 3d 3d 04 12 02 0d 00 00 64 00 00 00
00 00 00 00 00 05 03 00 00 60 06 09 01 00 00 00 00 00 00 00 00
```

Which in base64 becomes:

```
AoFuAE/////3/wN+AAFAYGRhdGE6YXBwbGljYXRpb24vbXBlZzQtb2QtYXU7YmFzZ
TY0LEFTb0JLQVVmQXlRQXlRQUVEU0FSQWdnQUFCYmpZQUFXNDJBR0VBQkVBQUFENk
FBQUFBQWdBQUFBQUFNPQQNAQUAAMgAAAAAAAAAAAYJAQAAAAAAAAAA2UAAkBCZGF
0YTphcHBsaWNhdGlvbi9tcGVnNC1iaWZzLWF1O2Jhc2U2NCx3QkFTWVFRZndBQUFI
OEFBQUVRb0lvS2ZnQT09BBICDQAAZAAAAAAAAAAABQMAAGAGCQEAAAAAAAAAAA==
```

12.3.7 ISMA1.0 audio-only IOD, profile 1

The binary representation of this IOD is identical to that of profile 0.