SMPTE STANDARD

for Television —
Declarative Data Essence — Internet Protocol Multicast Encapsulation

Table of contents

1 Scope and application

1.1 Scope

This standard defines the encapsulation of declarative data essence using internet protocol (IP) multicast. This is done in a transport-independent manner and relies solely on standard IP multicast techniques.

1.2 Application

This standard defines the transmission of declarative content across terrestrial (over the air), cable, and satellite systems as well as over the Internet. In addition, it will also bridge between networks — for example, data on an analog terrestrial broadcast must easily bridge to a digital cable system. This design goal was achieved through the definition of a transport-independent content format and the use of IP. Since IP bindings already exist for each of these video systems, the advantages of this work may be useful.

IP multicast is the mechanism for broadcast data delivery. Content creators should assume IP addresses may be changed downstream and, therefore, should not use them in their content. The transport operator is responsible only for making sure that an IP address is valid on the physical network where they broadcast it (not for any rebroadcasting). When possible, content creators should use valid IP multicast addresses to minimize the chance of collisions. Some systems may have two-way Internet connections. Capabilities in those systems are outside the scope of this standard and are described by the appropriate Internet standards.

Transport operators should use the standard IP transmission system for the appropriate medium (IETF, ATSC, DVB, etc.). It is assumed that when the user tunes to a television channel, the receiver automatically locates and delivers IP datagrams associated with the television broadcast. The mechanism for tuning video and connecting to the appropriate data stream is implementation and delivery standard specific and is not specified in this framework.

Approved
January 3, 2002

2 Normative references

The following standards contain provisions which, through reference in this text, constitute provisions of this standard. At the time of publication, the editions indicated were valid. All standards are subject to revision, and parties to agreements based on this standard are encouraged to investigate the possibility of applying the most recent edition of the standards indicated below.

SMPTE 363M-2002, Television — Declarative Data Essence — Content Level 1

SMPTE 364M-2001, Television — Declarative Data Essence — Unidirectional Hypertext Transport Protocol

IETF RFC 768, User Datagram Protocol (UDP)

IETF RFC 791, Internet Protocol — DARPA Internet Program Protocol Specification

IETF RFC 1112, Host Extensions for IP Multicasting

IETF RFC 2327, SDP: Session Description Protocol

IETF RFC 2974, Session Announcement Protocol (SAP)

ISO/IEC 11578:1996, Information Technology — Open Systems Interconnection — Remote Procedure Call (RPC), Annex A, Universal Unique Identifier (UUID)

3 Announcement protocol

Announcements are used to announce currently available programing to the receiver. The IP multicast addresses and ports for resource transfer and for triggers are announced using SDP announcements (IETF RFC 2327). The SDP header is preceded by an 8-byte SAP header (IETF RFC 2974). Announcements are sent on a well-known address (224.0.1.113) and port (2670). This address and port have been registered with the IANA.

v=0 SDP version, required to be 0.

o=username sid version IN IP4 ipaddress: Owner and session identifier, defined as usual in SDP specification. Username is "-", network type is IN, address type is IP4. *Sid* identifies an announcement for a particular broadcast (it can be a permanent announcement for all programming on a broadcast channel or for a particular show). *Version* indicates the version of the message. These values allow receivers to match a message to a previous message and know whether it has changed. Session ID and Version should be NTP values as recommended in SDP.

s=name Session name, required as in SDP specification.

i=, u= Optional, as in SDP specification.

e=, p= E-mail address or telephone number (at least one required in SDP specification).

b=CT:number Optional in SDP specification, but required here. Bandwidth in kilobytes per second as in the SDP specification. Bandwidth of the broadcast data can be used by receivers to choose among multiple versions of enhancement data according to the bandwidth the receiver can handle.

t=start stop As in SDP specification, gives start and stop time in NTP format. With programs stored on tape, at times it will not be possible to insert new announcements, so start times on tape could be incorrect. In this case, the start time should be set to the original broadcast time and the stop time set to 0. This is the standard for an

unboundedsession. Assumptions are then made about the stop time (see IETF RFC 2327). A new announcement with the same sid and different version for the same broadcast station replaces the previous one. It is preferred that a tool read the tape and generate announcements with correct start and stop times, but it is not required. Content creators can choose to use only a station ID and not provide information about individual programs.

a=UUID:UUID Optional. The UUID should uniquely identify the enhancement (for example, a different UUID for each program), and can be accessed using the trigger receiver object. In analog television and many types of digital television broadcast data are tied tightly to A/V. Each virtual channel has its own private network associated with it. In other systems, enhancements for many virtual channels can be carried on the same network. These systems can use the UUID to link a television broadcast with a particular enhancement. How that association is indicated is beyond the scope of this standard. One technique would be to place the UUID in electronic program guide information. Use ASCII hex to encode UUIDs.

a=type:tve Required. Indicates to the receiver that the announcement refers to an enhancement related to this specification.

a=lang, a=sdplang Optional, as in SDP specification.

a=tve-type:<types> Optional. *tve-type*: specifies an extensible list of types that describe the nature of the enhancement. It is a session-level attribute and is not dependent on *a=tve-type:primary*. Optional. *tve-type:primary* specifies that this will be the primary enhancement stream associated with the currently playing video program whenever this enhancement's trigger stream is active. If *tve-type:primary* is not specified, the TVE stream is never the primary enhancement stream associated with video. This, like all *tve-type:* attributes, is a session level attribute. This attribute can be used by receivers to implement automatic loading of primary video enhancement streams. The actual display of and switching between enhancement streams are handled by the trigger streams.

a=tve-size:Kbytes Required. tve-size: provides an estimate of the high watermark of cache storage in kilobytes that will be required during the playing of the enhancement. This is necessary so that receivers can adequately judge whether or not they can successfully play an enhancement from beginning to end.

a=tve-level:x Content level identifier, where x is 1.0 for this version of the framework (optional, default is 1.0).

a=tve-ends:seconds Optional, specifies an end time relative to the reception time of the SDP announcement. *Seconds* is the number of seconds in the future that this announcement is valid. *Seconds* may change (count down) as an announced session progresses. This attribute, when present, overrides the default assumptions for end times in unbounded announcements.

m=data portvalue/2 tve-file/tve-trigger
c=IN IP4 ipaddress/ttl As in SDP specification. Compact form specifying two ports on the same address.

When there are multiple alternative enhancement streams for the same video program, they must all be announced at the media level of the same SDP announcement. All enhancement streams announced in the same SDP announcement are considered to be mutually exclusive variants of the primary enhancement stream. The receiver can choose between them based on media level attributes. For example, the *a=lang* field can be used at the media level to choose between language variants of the primary enhancement.

Each media section for the tve-file media type begins the next enhancement definition.

A longer form is available if the content creator or transport operator wants to use different IP addresses and ports for the data stream and trigger stream:

m=data portvalue tve-file
c=IN IP4 ipaddress/ttl Alternative form for specifying addresses and ports (for file protocol, as in SDP specification).

m=data portvalue tve-trigger
c=IN IP4 ipaddress/ttl (for control protocol, as in SDP specification).

Announcement example:

v = 0
o = –2890844526 2890842807 IN IP4 tve.niceBroadcaster.com
s = Day & Night & Day Again
i = A very long TV Soap Opera
e = help@niceBroadcaster.com
a = UUID:f81d4fae-7dec-11d0-a765-00a0c91e6bf6
a = type:tve
a = tve-level:1.0
t = 2873397496 0
a = tve-ends:30000
a = tve-type:primary
m = data 52127/2 tve-file/tve-trigger
c = IN IP4 224.0.1.112/127
b = CT:100
a = tve-size:1024
m = data 52127/2 tve-file/tve-trigger
c = IN IP4 224.0.0.1/127
b = CT:1024
a = tve-size:4096

4 Trigger protocol

The trigger protocol carries a single trigger in a single UDP/IP multicast packet. Triggers are real-time events broadcast inside IP multicast packets delivered on the address and port defined in the SDP announcement for the enhanced television program (see announcements). The trigger protocol is thus very lightweight in order to provide quick synchronization.

5 Resource transfer: UHTTP

A one-way IP multicast based resource transfer protocol, the unidirectional hypertext transfer protocol (UHTTP), is defined in SMPTE 364M. UHTTP is a simple, robust, one-way resource transfer protocol that is designed to deliver efficiently resource data in a one-way broadcast-only environment. This resource transfer protocol is appropriate for IP multicast over television vertical blanking interval (IPVBI), in IP multicast carried in MPEG-2, or in other unidirectional transport systems.

Web pages and their related resources (such as images and scripts) are broadcast over UDP/IP multicast in the related television signal. An announcement broadcast by the television station tells the receiver which IP multicast address and port to listen to for the data. The only data broadcast to this address and port are resources intended for display as web content.

While HTTP headers preceding resource content are optional in the UHTTP protocol, they are required when the protocol is used for DDE. Compliant receivers must support *gzip* content encodings as specified by the content-encoding HTTP header field.

Annex A (informative)
Using enhanced television

Television enhancements are comprised of three related data sources: announcements (delivered via UDP using SAP/SDP), content (delivered via UHTTP), and triggers (delivered via the trigger protocol over UDP). Announcements are broadcast on a single well-known multicast address and have a time period for which they are valid. This time period is expressed via the "*t=*" and "*a=tve-ends:*" lines within the SDP record. Announcements also indicate the multicast address and port number that the client can listen in on to receive the content and triggers.

The announcement also contains information that the client optionally can use to help decide whether to automatically start receiving triggers and content information. This may include *a=tve-type, lang=*, and *keywds=* attributes that provide additional information to the client about the announced enhancements. For example, announcements with an optional *a=tve-type:primary* attribute may be used by the client to implement an auto-play feature. Multiple *a=tve-type* attributes may appear in a given announcement and are not mutually exclusive.

Annex B (informative)
Example broadcast

The following is a simple example of a television enhancement, delivered via transport type B (multicast IP). The example consists of three parts: the announcement (announced via SDP/SAP), the content (delivered via UHTTP), and the triggers (delivered in UDP packets).

The experience consists of a screen with 60% sized embedded live television object, with some text below it. During the show, a trigger may arrive that will cause an image of the word "MURDER" to appear below the text. If the user chooses to click on the television object, they will be returned to full- screen video, and away from the enhanced experience.

B.1 Announcement

The following announcement packet is sent via UDP to the multicast IP address: 224.0.1.113, port: 2670.

The announcement consists of an 8-byte SAP header followed by an SDP text payload.

The values for the SAP header fields for the announcement shall be as given in table B.1.

A complete SAP header would be 8 bytes: 0x20, 0x00, 0x34, 0x64, 0xd1, 0xf0, 0xc3, 0x06.

The remaining bytes in the announcement packet would contain the following text payload:

```
v=0
o=–2890844526 2890842807 IN IP4
     tve.niceBroadcaster.com
s=Day & Night & Day Again
i=A very long TV Soap Opera
e=help@niceBroadcaster.com
a=UUID:f81d4fae-7dec-11d0-a765-00a0c91e6bf6
a=type:tve
a=tve-level:1.0
a=tve-ends:1800
a=tve-type:primary
t=2873397496 0
m=data 52127/2 tve-file/tve-trigger
c=IN IP4 224.0.1.112/127
b=CT:40
a=tve-size:1024
```

These fields are indicated in table B.2.

Table B.1 – SAP header fields values

Field name (size)	Value	Description
Version (3 bits)	1	SAP version
Message type (3 bits)	0	Session description announcement packet
Encrypted (1 bit)	0	Not encrypted
Compressed (1 bit)	0	Not compressed
Authentication length (1 byte)	0x00	No authentication
Message ID hash (2 bytes)	0x3464	Hash of payload text
Originating source address (4 bytes)	209.240.195.6	IP address of originating host

Table B.2 – Description of fields

v=0	SDP version zero
o=–2890844526 2890842807 IN IP4 tve.niceBroadcaster.com	Originating host information
s=Day & Night & Day Again	Session name
i=n very long TV Soap Opera	Session description
e=help@niceBroadcaster.com	Contact information about the session
a=UUID:f81d4fae-7dec-11d0-a765-00a0c91e6bf6	Unique identifier (UUID) for the session
a=type:tve	This is a television enhancement
a=tve-level:1.0	Content level 1
a=tve-ends:1800	Session ends 30 minutes from now
a=tve-type:primary	This session is the primary enhancement to the video
t=2873397496 0	Session began at a particular time
m=data 52127/2 tve-file/tve-trigger	File and trigger data are available on ports 52127 and 52127+1
c=IN IP4 224.0.1.112/127	Data will be broadcast on multicast address 224.0.1.112
b=CT:40	This session will have a maximum band-width of 40 kb/s
a=tve-size:3	This session will require a maximum amount of caching of 3k bytes

B.2 Content

The content data for the enhancement is delivered via UHTTP packets transmitted (as specified by the announcement) to multicast address 224.0.1.112 and to port 52127.

This content would consist of three original source files, two HTML documents, and a PNG image. The experience consists of a screen with a 60% sized embedded live television object, with some text below it. During the show, a trigger may arrive that will cause an image of the word "MURDER" to appear below the text. If the user chooses to click on the television object, they will be returned to full screen video, and away from the enhanced experience.

The first would be referred to by the URL:

<lid://nicebroadcaster.com/show27/launch.html>, and consists of the following text:

```
<HTML>

<OBJECT TYPE="application/tve-trigger" ID= "triggerReceiverObj>"
</OBJECT>

<HEAD>
<TITLE>Day & Night & Day: The Interactive Experience</TITLE>
</HEAD>

<BODY bgcolor="magenta>"

<A href="tv:">
<OBJECT data="tv:" width="60%" height="60%" align="center">
</OBJECT>
</A>

<BR>
<P>Welcome to the Day & Night & Day Interactive Experience</P><BR>
<P>Watch below for more information about the current scene!</P><BR>

<IMG name="sceneimage" align="center" src="">

</BODY>
</HTML>
```

The other files consist of a second HTML file and a PNG image, containing an image of the word "MURDER" in big red letters. Its URL will be specified as <lid://nicebroadcaster.com/show27/murder.png>

These files are combined together into a single multipart MIME entity, which will make up the full payload of the UHTTP transmission.

```
Content-Base: lid://nicebroadcaster.com/show27
Content-Length: 2264
Content-Type: Multipart/Related; boundary=example98203804805

–example98203804805
Content-Location: launch.html
Content-Length: 450
Content-Type: text/html

<HTML>

</OBJECT TYPE="application/tve-trigger" ID="triggerReceiverObj">
</OBJECT>

<HEAD>
<TITLE>Day & Night & Day: The Interactive Experience</TITLE>
</HEAD>

<BODY bgcolor="magenta">

<A href="tv:">
<OBJECT data="tv:" width="60%" height="60%" align="center">
</OBJECT>
</A>

<BR>
<P>Welcome to the Day & Night & Day Interactive Experience</P><BR>
<P>Watch below for more information about the current scene!</P><BR>

</BODY>
</HTML>
–example98203804805
Content-Location: murder.html
Content-Length: 475
Content-Type: text/html
```

```
<HTML>

<OBJECT TYPE="application/tve-trigger" ID="triggerReceiverObj">
</OBJECT>

<HEAD>
<TITLE>Day & Night & Day:  The Interactive Experience</TITLE>
</HEAD>

<BODY bgcolor="magenta">

<A href="tv:">
<OBJECT data="tv:" width="60%" height="60%" align="center">
</OBJECT>
</A>

<BR>
<P>Welcome to the Day & Night & Day Interactive Experience</P><BR>
<P>Watch below for more information about the current scene!</P><BR>

<IMG align="center" src="murder.png">

</BODY>
</HTML>

–example98203804805
Content-Location:  murder.png
Content-Length:  1289
Content-Type:  image/png

binary resource data for murder.png image
–example98203804805
```

This data multipart entity (total length of 2400 bytes, including headers) would be transmitted via UHTTP in three packets. The first two packets would contain the original data (each containing 1200 bytes of original data as payload) and the third containing the exclusive-or of the first and second payloads as forward error correction data.

The UHTTP headers for each of the three packets shall be as given in table B.3.

Table B.3 – Headers for packets

Field (size)	Packet 1 value	Packet 2 value	Packet 3 value	Description
Version (5 bits)	0000	0000	0000	UHTTP version
ExtensionHeader (1 bit)	0	0	0	No extension headers in this example
HTTPHeadersPrecede (1 bit)	1	1	1	HTTP headers precede data
CRCFollows (1 bit)	0	0	0	No CRC follows
PacketsInXORBlock (1 byte)	3	3	3	Number of packets in each XOR FEC block
Retransmit Expiration (2 bytes)	1800	1800	1800	This will be retransmitted for the next 1800 seconds (this value will decrease as the show progresses)
TransferID (16 bytes)	0x1432 3ab4123 ab4567c d89ef05 67cd89e f0	same	same	UUID for this transmission
Resource Size (4 bytes)	2400	2400	2400	Size of payload
SegStartByte (4 bytes)	0	1200	2400	Offset into the stream where this packet's payload starts

In this example, the 28 UHTTP header bytes for the first packet would be:

0x02, (version, options)
0x03, (packets in XOR block)
0x07, 0x08, (retransmit expiration)
0x14, 0x32, 0x3a, 0xb4, 0x12, 0x3a, 0xb4, 0x56, 0x7c, 0xd8, 0x9e, 0xf0, 0x56,
0x7c, 0xd8, 0x9e, 0xf0, (Resource ID)
0x00, 0x00, 0x09, 0x2e, (Resource Size)
0x00, 0x00, 0x00 (Segment Start Byte Offset)

Following the header in the first packet would be the first 1200 bytes of the MIME-encoded payload. Following the header in the second packet would be the last 1175 bytes of the MIME encoded payload. Following the UHTTP header in the third packet would be 1200 bytes, where each byte is the exclusive-or of the corresponding byte offsets in the first two packets.

These packets would then be transmitted repeatedly during the session. The header values for each packet would remain the same, with the exception of the Retransmit Expiration field. The value of this field would decrease as the end of the transmission of the UHTTP packets drew near.

B.3 Triggers

The following trigger would be sent after the data are first transmitted to trigger the beginning of the enhanced television experience:

<lid://nicebroadcaster.com/show27/launch.html>[name:Day & Night & Day Again Interactive]

This trigger content would be encapsulated in a UDP packet and sent to multicast address: 224.0.1.112, port 52127+1 (as specified by the announcement).

This trigger packet would also be transmitted periodically later on to allow viewers who tune in late to join in the experience.

Later on, during the program, the content creator might send the following trigger to the same multicast address and port to make the content change to reflect the fact that a murder scene has just begun in the program:

<lid://nicebroadcaster.com/show27/launch.html>[s:window.top.location.href="lid://nicebroadcaster.com/show27/murder.html"]

This trigger would cause the active enhancement page (if it matched the URL in the trigger) to load the new page, "murder.html", which would cause the murder.png image to be displayed within the page. If the specified URL was not currently being displayed, the trigger would be ignored because this trigger does not include a [name:] attribute.

Near the end of their program, they might send the following trigger to tell their interactive application to shut down. This would allow them to synchronize more accurately with the end of the program, rather than relying on the session timing information in the announcement.

<lid://nicebroadcaster.com/show27/launch.html>[script: window.location="tv:"]

Annex C (informative)
Glossary

C.1 announcements: Announcements are used to announce currently available programing to the receiver.

C.2 binding: In the context of this standard, a binding is the definition of how the transport specifications are encoded on a specific video network standard.

C.3 content creator: In the context of this standard, a content creator has the role of originating the content components of the television enhancement including graphics, layout, interaction, and triggers.

C.4 datagram: A block of data that is smart enough (actually, which carries enough information) to travel from one Internet site to another without having to rely on earlier exchanges between the source and destination computer.

C.5 essence: Raw program material.

C.6 FEC: Forward error correction.

C.7 FTP (file transfer protocol): A standard for finding and transferring files on the Internet.

C.8 HTTP (hypertext transfer protocol): A set of instructions for communication between a server and a World Wide Web client.

C.9 IANA (Internet Assigned Numbers Authority): The central registry for various Internet protocol parameters, such as port, protocol and enterprise numbers, and options, codes, and types. The currently assigned values are listed in the assigned numbers document.

C.10 IETF (Internet Engineering Task Force): The IETF is a large, open community of network designers, operators, vendors, and researchers whose purpose is to coordinate the operation, management, and evolution of the Internet, and to resolve short-range and midrange protocol and architectural issues. It is a major source of proposals for protocol standards which are submitted to the IAB for final approval. The IETF meets three times a year and extensive minutes are included in the IETF Proceedings.

C.11 IP (Internet protocol): This protocol is one of the languages computers connected to the Internet use to communicate.

C.12 IP multicast: A one-to-many transmission, in contrast to unicast broadcast. An extension to the standard IP network-level protocol. IETF RFC 1112, host extensions for IP multicasting, authored by Steve Deering in 1989, laid the groundwork for IP multicasting. The RFC describes IP multi-casting as the transmission of an IP datagram to a host group, a set of zero or more hosts identified by a single IP destination address. A multicast datagram is delivered to all members of its destination host group with the same best- efforts reliability as regular unicast IP datagrams. The membership of a host group is dynamic; that is, hosts may join and leave groups at any time. There is no restriction on the location or number of members in a host group. A host may be a member of more than one group at a time.

C.13 ISO (International Organization for Standardization): A voluntary, nontreaty organization founded in 1946 which is responsible for creating international standards in many areas including computers and communications. Its members are the national standards organizations of the 89 member countries, including ANSI for the U.S.

C.14 MIME (multipurpose Internet mail extensions): A protocol for allowing e-mail messages to contain various types of media (text, audio, video, images, etc.).

C.15 NABTS: North American Basic Teletext Specification.

C.16 receiver: In the context of this standard, a receiver is a hardware and software implementation (television, set-top box, or personal computer) that decodes and presents enhanced content.

C.17 SAP (session announcement protocol): The protocol used for session announcements.

C.18 SDP (session description protocol): SDP is intended for describing multimedia sessions for the purposes of session announcement, session invitation, and other forms of multimedia session initiation.

C.19 transport operator: In the context of this standard, the transport operator runs a video delivery infrastructure (terrestrial, cable, satellite, or other) that includes a transport for enhancement data.

C.20 triggers: A text message used to announce the availability of an enhancement to the receiver, or execute ECMA- Script on a document in a running enhancement. Triggers include the URL of the enhancement and optionally a human-readable string to use in order to announce the availability of the interactive television experience to the user.

C.21 television enhancement: A collection of web content displayed in conjunction with a television broadcast as an enhanced or interactive program.

C.22 UDP (user datagram protocol): An Internet standard transport layer protocol defined in STD 6, IETF RFC 768. It is a connectionless protocol which adds a level of reliability and multiplexing to IP.

C.23 UHTTP (unidirectional hypertext transfer protocol): UHTTP is a simple, robust, one-way resource transfer protocol that is designed to deliver efficiently resource data in a one-way broadcast-only environment. This resource transfer protocol is appropriate for IP multicast over television vertical blanking interval (IPVBI), in IP multicast carried in MPEG-2, or in other unidirectional transport systems.

C.24 UUID (universally unique identifier): Also known as GUID (globally unique identifier) is an identifier that is unique across both space and time, with respect to the space of all UUIDs.