

CableLabs® Specifications Web Technology

Mapping from MPEG-2 Transport to HTML5

CL-SP-HTML5-MAP-I01-120120

ISSUED

Notice

This CableLabs specification is the result of a cooperative effort undertaken at the direction of Cable Television Laboratories, Inc. for the benefit of the cable industry and its customers. This document may contain references to other documents not owned or controlled by CableLabs. Use and understanding of this document may require access to such other documents. Designing, manufacturing, distributing, using, selling, or servicing products, or providing services, based on this document may require intellectual property licenses from third parties for technology referenced in this document.

Neither CableLabs nor any member company is responsible to any party for any liability of any nature whatsoever resulting from or arising out of use or reliance upon this document, or any document referenced herein. This document is furnished on an "AS IS" basis and neither CableLabs nor its members provides any representation or warranty, express or implied, regarding the accuracy, completeness, noninfringement, or fitness for a particular purpose of this document, or any document referenced herein.

© 2012 Cable Television Laboratories, Inc.
All rights reserved.

Document Status Sheet

Document Control Number:	CL-SP-HTML5-MAP-I01-120120			
Document Title:	Mapping from MPEG-2 Transport to HTML5			
Revision History:	I01 - Released 01/20/12			
Date:	January 20, 2012			
Status:	Work in Progress	Draft	Issued	Closed
Distribution Restrictions:	Author Only	CL/Member	CL/Member/Vendor	Public

Key to Document Status Codes

Work in Progress	An incomplete document, designed to guide discussion and generate feedback that may include several alternative requirements for consideration.
Draft	A document in specification format considered largely complete, but lacking review by Members and vendors. Drafts are susceptible to substantial change during the review process.
Issued	A stable document, which has undergone rigorous member and vendor review and is suitable for product design and development, cross-vendor interoperability, and for certification testing.
Closed	A static document, reviewed, tested, validated, and closed to further engineering change requests to the specification through CableLabs.

Trademarks

CableCARD™, CableHome®, CableLabs®, CableNET®, CableOffice™, CablePC™, DCAS™, DOCSIS®, DPoE™, EBIF™, eDOCSIS™, EuroDOCSIS™, EuroPacketCable™, Go2BroadbandSM, M-Card™, M-CMTS™, OCAP™, OpenCable™, PacketCable™, PCMM™, PeerConnect™, and tru2way® are marks of Cable Television Laboratories, Inc. All other marks are the property of their respective owners.

Table of Contents

- 1 SCOPE.....1**
- 1.1 Introduction and Purpose 1
- 1.2 Requirements 2
- 2 REFERENCES3**
- 2.1 Normative References.....3
- 2.2 Informative References3
- 3 TERMS AND DEFINITIONS4**
- 4 ABBREVIATIONS AND ACRONYMS.....5**
- 5 REQUIREMENTS6**
- 5.1 Video, Audio and Text Track Creation.....6
 - 5.1.1 *Program Description TextTrack*.....6
 - 5.1.2 *VideoTrack*.....7
 - 5.1.3 *AudioTrack*7
 - 5.1.4 *Other TextTracks*7
- 5.2 Closed Captioning8

Figures

- Figure 1 - Sourcing In-band Media Resource Tracks 1

This page left blank intentionally.

1 SCOPE

1.1 Introduction and Purpose

HTML5 user agents (UAs), per [HTML5], may playback MPEG-2 TS media resources that contain a multiplex of video, audio, text, and private data elementary streams. Television program providers and distributors use these streams to deliver services associated with the primary video and audio in the multiplexed stream. These services are collectively termed “TV Services”. A common HTML representation of these TV services tracks is essential in order for these TV Services to be made available to Web content in a consistent way. This specification defines requirements for how these MPEG-2 TS elementary streams should be translated by the HTML5 UA into the equivalent HTML5 video, audio, and text track elements. Figure 1 illustrates the relationship of this translation function in the context of MPEG-2 TS media resources delivered to a Web page by a UA in a client application.

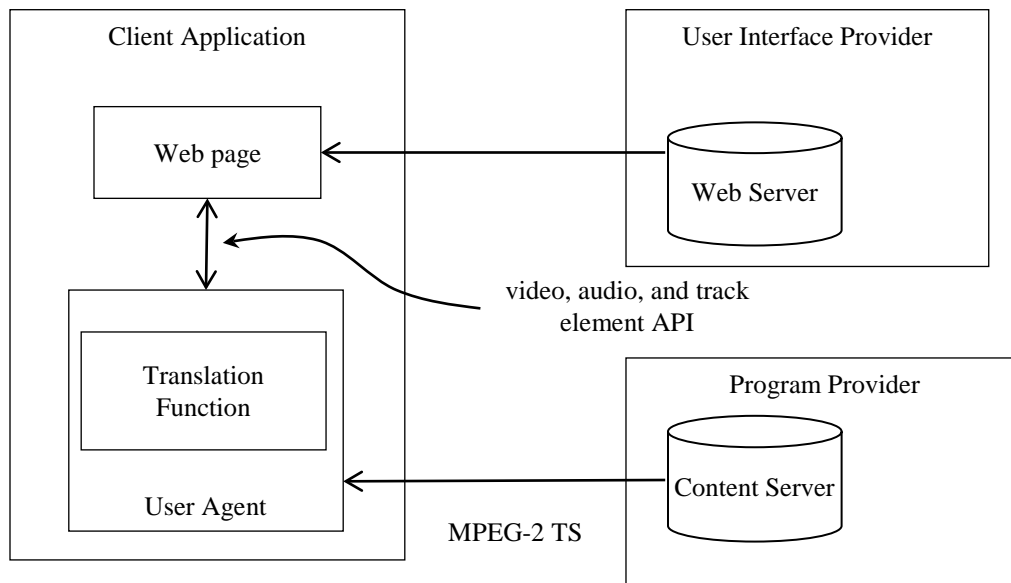


Figure 1 - Sourcing In-band Media Resource Tracks

Note that the Web page providing the user interface (e.g., program guide) is often not provided by the originator of the program content. For example, the guide may be provided by the television manufacturer or the cable or satellite TV provider, while the multiplexed streams are provided by hundreds of independent television program providers. Therefore, the Web page has no *a priori* knowledge of which streams are in the programs at any given time.

1.2 Requirements

Throughout this document, the words that are used to define the significance of particular requirements are capitalized. These words are:

- | | |
|--------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| "MUST" | This word means that the item is an absolute requirement of this specification. |
| "MUST NOT" | This phrase means that the item is an absolute prohibition of this specification. |
| "SHOULD" | This word means that there may exist valid reasons in particular circumstances to ignore this item, but the full implications should be understood and the case carefully weighed before choosing a different course. |
| "SHOULD NOT" | This phrase means that there may exist valid reasons in particular circumstances when the listed behavior is acceptable or even useful, but the full implications should be understood and the case carefully weighed before implementing any behavior described with this label. |
| "MAY" | This word means that this item is truly optional. One vendor may choose to include the item because a particular marketplace requires it or because it enhances the product, for example; another vendor may omit the same item. |

2 REFERENCES

2.1 Normative References

In order to claim compliance with this specification, it is necessary to conform to the following standards and other works as indicated, in addition to the other requirements of this specification. Notwithstanding, intellectual property rights may be required to use or implement such normative references.

- [HTML5] “HTML5 - A vocabulary and associated APIs for HTML and XHTML”, <http://www.w3.org/TR/html5/>
- [RFC 3555] IETF RFC 3555, “MIME Type Registration of RTP Payload Formats”, <http://tools.ietf.org/html/rfc3555>
- [ISO 639.2] ISO 639.2, "Code for the Representation of Names of Languages — Part 2: alpha-3 code", www.loc.gov/standards/iso639-2/iso639jac.html
- [RFC 4648] IETF RFC 4648, “The Base16, Base32, and Base64 Data Encodings”, <http://www.ietf.org/rfc/rfc4648.txt>
- [H.222.0] ISO/IEC 13818-1|ITU-T H.222.0 5/2006, “Infrastructure of audiovisual services – Transmission multiplexing and synchronization”, <http://www.itu.int/rec/T-REC-h.222.0/en>
- [BCP 47] IETF BCP 47, “Tags for Identifying Languages”, <http://www.ietf.org/rfc/bcp/bcp47.txt>

2.2 Informative References

- [JSON] “JSON in JavaScript”, <http://www.json.org/js.html>

3 TERMS AND DEFINITIONS

This specification uses the following terms:

Descriptor	Structure, used to extend the definitions of programs and streams, consisting of an 8-bit tag followed by an 8-bit descriptor length and data fields.
Elementary Stream	A generic term for a coded video, audio, or other data stream carried in a sequence of PES packets with the same stream id.
Media Resource Timeline	Maps times (in seconds) to positions in the media resource. The origin of a timeline is its earliest defined position. The duration of a timeline is its last defined position. Identical to the media timeline defined in [HTML5].
Packet Identifier	A number that uniquely identifies an elementary stream in a program.
Packetized Elementary Stream	An elementary stream encoded in sequence of PES packets where each packet consists of a header followed by a number of contiguous bytes from the elementary stream.
Program Map Table	Specifies the PID stream type, PID, and optional stream descriptors that identify the elementary streams that form each program.
Program Stream	A generic term for an elementary stream that is part of a program.
User Agent	The function that conforms to the HTML specification as defined in [HTML5]. A Web browser is a common example of a user agent.

4 ABBREVIATIONS AND ACRONYMS

This specification uses the following abbreviations:

DASH	Dynamic adaptive streaming over HTTP
HTML	Hypertext markup language
MPEG-2 TS	MPEG-2 transport stream
PES	Packetized elementary stream
PID	Packer identifier
PMT	Program map table
UA	user agent

5 REQUIREMENTS

This specification defines the requirements for an HTML5 user agent (UA) to recognize and make available to Web content the MPEG-2 TS program streams so that the following set of TV Services can be provided:

Closed Captioning	Textual representation of the media resource audio dialogue intended for the hearing impaired.
Subtitles	Alternate language textual representation of the media resource audio dialogue.
Content Advisories	Content rating information used by parental control applications.
Synchronized Content	Signaling messages to control the execution of a client application in a manner synchronized with the media resource playback.
Client ad insertion	Signaling messages that convey advertisement insertion opportunities to a client application.
Audio translations	Alternate language representation of the primary audio track.
Audio descriptions	Audio descriptions of the video intended for the visually impaired.

The requirements in this specification only apply to single program MPEG-2 transport streams; multi-program MPEG-2 transport streams are out of scope for this specification.

Other specifications may define equivalent requirements for media transport and container formats other than MPEG-2 TS, MPEG-4 base media file format, and MPEG DASH, for example.

The following sub-sections define requirements for how the UA is required to recognize MPEG-2 TS video, audio, and other data tracks, and how the HTML5 elements representing those tracks are to be created.

HTML5 VideoTrack, AudioTrack and TextTrack elements have additional attributes, beyond those referenced in this specification that SHOULD be set by the UA, consistent with user preferences.

A UA may be presented with previously processed in-band TV services tracks, for example, when the viewer seeks back in the media resource, as controlled by the seekable time ranges attribute of the HTMLMediaElement. Text track cues are not removed from the text track, so the UA MUST NOT create duplicate text track cues in this case. How the UA accomplishes this is implementation-specific.

5.1 Video, Audio and Text Track Creation

HTML5 VideoTrack, AudioTrack, and TextTrack MUST be created as defined in [HTML5].

5.1.1 Program Description TextTrack

Recognition of specific types of video, audio, and text tracks will, in general, be dependent on geographical region or service or content provider. In order that UA implementations are independent of region and provider, it is desirable that the UA recognize tracks in a generic manner and rely on a script to implement region, provider, and application specific recognition of tracks. This requires script access to the MPEG program description in the program map table (PMT) as defined in [H.222.0] so that the tracks can be used correctly.

The UA MUST create a TextTrack in the media resource TextTrack list to make the PMT available to a script and set the TextTrack attributes using the following rules:

- 1) kind = "metadata"
- 2) label = "video/mp2t track-description", i.e., the string concatenation of the MPEG-2 TS MIME type [RFC 3555] and " track-description"
- 3) language = "" (empty string)
- 4) mode= TextTrack HIDDEN

For each PMT received in the program stream, the UA MUST create a TextTrackCue only in the case where the PMT differs from the PMT represented by the previously created TextTrackCue. This is in recognition of the fact that the PMT is received at a minimum rate of every 140 msec but rarely changes.

For each new PMT, a UA MUST create a new TextTrackCue in the text track as described in [HTML5] section “Text track model” with attributes set as follows:

- 1) `startTime` is set to the current time in the media resource timeline
- 2) `endTime` = Infinity
- 3) `text` is set to a JSON [JSON] string representation of the PMT:
`'{"m2pt_track_description":[program_info_entry, ...]}'`. Each *program_info_entry* is `{"stream_type":string representation of the PMT stream_type field, "pid":string representation of the PMT elementary_PID field, "es_descriptors":[es_descriptor, ...]}`. Each *es_descriptor* is `{"tag":string representation of the PMT descriptor tag, "desc_contents":Base64 encoded representation of the PMT elementary stream descriptor}` (see [RFC 4648]).
- 4) `pauseOnExit`=false

`TextTrackCue.getCueAsHTML()` MUST return `""` (empty string)

5.1.2 VideoTrack

For all MPEG-2 video stream types supported by the UA, the UA MUST create a new VideoTrack in the VideoTrackList of the media resource.

The HTML5 specification, [HTML5], requires the VideoTrackList use the order defined by the media resource. VideoTracks MUST appear in the VideoTrackList in the same order as they appear in the PMT as defined in [H.222.0].

The UA MUST set the VideoTrack label attribute to the text string "video/mp2t-*pid*" where *pid* is the PID in the PMT for that track.

The UA MUST set `VideoTrackList[0].VideoTrack.kind` = “main”.

For all VideoTrackList entries, the UA SHOULD set the VideoTrack kind and language attributes if it can determine the correct values.

If the UA cannot determine the VideoTrack kind and language attributes it MUST set them to the empty string.

5.1.3 AudioTrack

For all MPEG-2 audio stream types supported by the UA, the UA MUST create a new AudioTrack in the AudioTrackList of the media resource.

The [HTML5] specification requires the AudioTrackList use the order defined by the media resource. AudioTracks MUST appear in the AudioTrackList in the same order as they appear in the PMT as defined in [H.222.0].

For each AudioTrack created, the UA MUST set the label attribute to the text string "video/mp2t-*pid*" where *pid* is the PID in the PMT for that track.

The UA MUST set `AudioTrackList[0].AudioTrack.kind` = “main”.

For all AudioTrackList entries, the UA SHOULD set the AudioTrack kind and language attributes if it can determine the correct values.

If the UA cannot determine the AudioTrack kind and language attributes, it MUST set them to the empty string.

5.1.4 Other TextTracks

For all MPEG-2 stream types that are not audio or video stream types, the UA MUST create a new TextTrack in the TextTrackList of the media resource.

The [HTML5] specification requires the TextTrackList use the order defined by the media resource. TextTracks MUST appear in the TextTrackList in the same order as they appear in the PMT as defined in [H.222.0].

The UA SHOULD set the TextTrack kind attribute to one of the categories defined in [HTML5].

If the UA cannot determine the TextTrack kind attribute, it MUST set it to “metadata”.

If the UA sets the TextTrack.kind attribute to one of the categories defined in [HTML5], it SHOULD set the TextTrack language attribute if it can determine the appropriate value.

If the UA cannot determine the TextTrack language attribute, it MUST set it to the empty string.

The UA MUST:

1. set the label attribute to the text string "video/mp2t-pid" where *pid* is the PID in the PMT for that track
2. set mode= TextTrack DISABLED

The MPEG-2 TS packets with the PID corresponding to the TextTrack contain either PES packets or private data packets as defined in [H.222.0]. For each PES or private data packet in the program stream represented by the TextTrack, the UA MUST create a TextTrackCue in the TextTrack as described in [HTML5] section “Text track model” with attributes set as follows:

- 1) startTime is set to the current time in the media resource timeline
- 2) endTime is set to the Infinity
- 3) text is set to the PES or private data packet encoded in Base64, see [RFC 4648].
- 4) pauseOnExit is set to false

TextTrackCue.getCueAsHTML() MUST return "" (empty string)

It is important to note that the above metadata TextTrack and TextTrackCue creation requirements, while minimizing region/provider/application knowledge in the UA, make the semantics of the metadata TextTrack and TextTrackCues opaque to the UA. So, for example, if the UA does not recognize subtitle tracks but creates a generic metadata text track as defined above, the UA behavior defined in [HTML5] for subtitle tracks will not occur since the UA is not aware this is a subtitle track. It is up to a script to identify the subtitle track and process the subtitle messages in the TextTrackCues in a manner appropriate for the subtitle format. One way this could be done is for a script to receive the TextTrackCues, extract the subtitle messages, and create a new subtitle TextTrack and TextTrackCues, in which case the UA defined behavior in [HTML5] would occur.

5.2 Closed Captioning

Closed captioning is delivered as part of the MPEG-2 TS video stream and MUST be recognized and made available by the UA.

A UA that recognizes closed captioning MUST:

- 1) Create a new TextTrack as defined in [HTML5] section “Sourcing in-band text tracks” with the track element attributes set as follows:
 - a. kind=“caption”
 - b. language is set to a [BCP 47]-conformant representation of the caption data language
 - c. label is set to a text string of the decimal representation of the PID of the MPEG-2 video program stream containing the caption data
- 2) Create a new TextTrackCue in the TextTrack as described in [HTML5] for each caption with attributes set as follows:
 - a. startTime is set to the presentation time set in the caption data converted into the equivalent time in seconds relative to the media resource timeline
 - b. endTime is set to the end of the presentation if specified in the caption data. If the end of the presentation is not specified, endTime is set to Infinity.

- c. text is set to the caption data if the UA determines that the caption data is in a plain text format compatible, or it can make compatible, with a UTF-16 string representation. Otherwise, the text attribute is set to a Base64 representation of the binary octet string composing the caption data. It is UA implementation-specific as to how the type of the caption data cue is determined by the UA when the cue text is rendered.
 - d. pauseOnExit is set to false.
- 3) Return a DocumentFragment with an HTML representation of the TextTrackCue text attribute as defined in [HTML5], if the UA knows how to create such a representation. Otherwise, return null.

