ETSI TS 103 286-2 V1.2.1 (2017-08)

Digital Video Broadcasting (DVB);
Companion Screens and Streams;
Part 2: Content Identification and Media Synchronization

EBU
OPERATING EUROVISION

DVB
Digital Video Broadcasting

Reference
RTS/JTC-DVB-374-2

Keywords
broadcast, companion screen, DVB,
synchronization, TV

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - NAF 742 C
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° 7803/88

Contents

Intellectual Property Rights

Essential patents

IPRs essential or potentially essential to the present document may have been declared to ETSI. The information pertaining to these essential IPRs, if any, is publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (https://ipr.etsi.org/).

Pursuant to the ETSI IPR Policy, no investigation, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Trademarks

The present document may include trademarks and/or tradenames which are asserted and/or registered by their owners. ETSI claims no ownership of these except for any which are indicated as being the property of ETSI, and conveys no right to use or reproduce any trademark and/or tradename. Mention of those trademarks in the present document does not constitute an endorsement by ETSI of products, services or organizations associated with those trademarks.

Foreword

This Technical Specification (TS) has been produced by Joint Technical Committee (JTC) Broadcast of the European Broadcasting Union (EBU), Comité Européen de Normalisation ELECtrotechnique (CENELEC) and the European Telecommunications Standards Institute (ETSI).

NOTE: The EBU/ETSI JTC Broadcast was established in 1990 to co-ordinate the drafting of standards in the specific field of broadcasting and related fields. Since 1995 the JTC Broadcast became a tripartite body by including in the Memorandum of Understanding also CENELEC, which is responsible for the standardization of radio and television receivers. The EBU is a professional association of broadcasting organizations whose work includes the co-ordination of its members' activities in the technical, legal, programme-making and programme-exchange domains. The EBU has active members in about 60 countries in the European broadcasting area; its headquarters is in Geneva.

European Broadcasting Union
CH-1218 GRAND SACONNEX (Geneva)
Switzerland
Tel: +41 22 717 21 11
Fax: +41 22 717 24 81

The Digital Video Broadcasting Project (DVB) is an industry-led consortium of broadcasters, manufacturers, network operators, software developers, regulatory bodies, content owners and others committed to designing global standards for the delivery of digital television and data services. DVB fosters market driven solutions that meet the needs and economic circumstances of broadcast industry stakeholders and consumers. DVB standards cover all aspects of digital television from transmission through interfacing, conditional access and interactivity for digital video, audio and data. The consortium came together in 1993 to provide global standardization, interoperability and future proof specifications.

The present document is part 2 of a multi-part deliverable covering the DVB Companion Screens and Streams Specification, as identified below:

Part 1: "Concepts, roles and overall architecture";

Part 2: "Content Identification and Media Synchronization";

Part 3: "Discovery".

Modal verbs terminology

In the present document "**shall**", "**shall not**", "**should**", "**should not**", "**may**", "**need not**", "**will**", "**will not**", "**can**" and "**cannot**" are to be interpreted as described in clause 3.2 of the ETSI Drafting Rules (Verbal forms for the expression of provisions).

"**must**" and "**must not**" are **NOT** allowed in ETSI deliverables except when used in direct citation.

Introduction

It is assumed that the reader is familiar with ETSI TS 103 286-1 [i.10] that provides background information on the concepts covered in the present document.

Personal, smart devices like tablet computers and smart phones enable new user experiences for broadcast service consumption. Many of these new experiences require synchronization between the broadcast content presented on the TV Device and the content presented on the personal device.

The present document enables the identification of, and synchronization with timed content and trigger events on TV devices (for example a Connected TV or STB) and related content presented by an application running on a personal device. Example use cases are:

- presenting a question and a choice of possible answers on a personal device, which are related to what is currently happening on a quiz show that is the current TV Programme;
- alternative audio intended to be consumed through the personal device (e.g. via connected headphones) and along with the broadcast video content on the TV device, such as an alternative commentary, an alternative language, clean audio for hearing impaired or audio descriptions for the visually impaired;
- seamlessly switching between different camera views on the personal device synchronously with a football game on the TV Device;
- presenting advertisements on the personal device which are related to the broadcast content, and in synchronization with the broadcast content (for example an advertisement for a product that is currently prominently visible in the broadcast video); and
- presenting a targeted advertisement to the user on the personal device at the time of presenting a generic interstitial in the broadcast content.

To enable such use cases, two functionalities are needed. The first functionality is the identification of broadcast content and finding of associated content for presentation on the personal device. The identification mechanisms defined in the present document are hence designed to take the following limiting factors into account:

- Different broadcast platforms may use different solutions to identify broadcast programmes.
- Synchronized transmission of broadcast and personal content through different transmission infrastructures is difficult.
- Broadcast platforms are bandwidth sensitive, and the amount of data needed for identification and synchronization should hence be kept to a minimum.
- Interactions of the personal device or the TV device with resources over broadband connections may take significant time. Furthermore, in a series of requests, the amount of time spent on each of them can vary largely and unpredictably.
- Applications presenting content on the personal device, and which are controlled by trigger events, need to identify of the content currently being presented by the TV device and determine the associated content for presentation on the personal device. In addition they also will need to subscribe to, and receive notifications of, any trigger event signalling in the broadcast service.

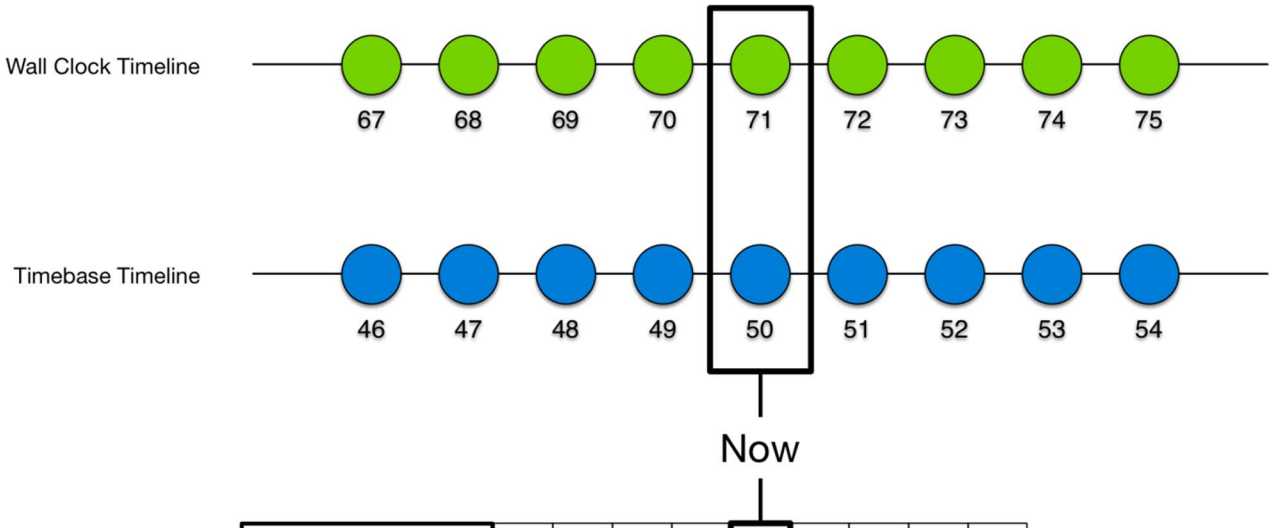
The second functionality is the synchronization of playback on the personal device with the playback on the TV device. Typically, an accuracy of at least 40 ms is required for frame-accurate synchronization between two video streams or lip sync between audio and video. The synchronicity between pieces of content is influenced by several factors:

- Propagation delays are different for different transmission networks and technologies, like terrestrial broadcast, satellite broadcast, IP multicast, and IP unicast; this can lead to arrival time differences of up to ten secs when transmitting through different paths.
- If the media is delivered via a Content Delivery Network (CDN), a significant amount of time (often 30 secs or more) is needed to ingest the content into the CDN before it becomes available for retrieval.
- Media processing function like transcoding can take up to several seconds of time which may limit their applicability to certain use cases.
- Streaming delivery through switched packet networks (for instance typical broadband Internet services) requires buffering for smooth presentation of media content. The size of the buffer depends on the technology used and the presence of any local post-processing for quality enhancement. This may lead to additional delays of up to 500 ms at the receiving device.

The present document provides an architectural framework for timeline synchronization between a presentation on one device and associated Timed Content on another, optionally using trigger events. The present document provides the protocol interfaces to provide this functionality given the limiting and influencing factors described above. These mechanisms are designed to take the following additional limiting factors into account:

- Related pieces of content may have different types of Timelines, with different tick rates and resolutions.
- The clocks of related pieces of content may exhibit different error properties (e.g. drift or jitter) if these clocks are not genlocked ("generator locked", i.e. synchronized at the source).
- As a consequence of processing during content production or distribution, timelines may be offset between different related pieces of content, even if they are of the same type and if clocks are genlocked.
- Timelines embedded into content (e.g. PTS for MPEG TS, or composition time of ISOBMFF) may be modified by the distribution network due to multiplexing, transcoding, and re-origination.
- Timelines transmitted along with content may be removed by distribution networks.
- Timelines can, and will, include discontinuities.
- Depending on the particular encoding of timestamps, some timelines will wrap around during presentation, as only a limited number of bits is available to express a Time Value on such Timeline.
- The system clocks of TV devices and personal devices run independently, and will hence exhibit different error behaviour (e.g. drift).

Figure 0.1 shows a basic, conceptual model for time-controlled playback. A local wall clock advances steadily, and the playback of the media streams is timed accordingly to achieve a smooth presentation. To enable this, the media streams are adorned with their own timebase timeline, which is compared to the wall clock timeline. During playback, whenever the wall clock timeline advances, the media player computes the corresponding point on the media's timebase timeline, and retrieves the associated chunk of media data for playback. To achieve time-controlled and smooth playback the media player will typically apply an offset to the media stream's timebase timeline and also adjust the playback rate of the media stream's timebase timeline in these computations. As the playback proceeds over time, media players will typically make dynamic re-adjustments of the offset and playback rate, to accommodate variations in the wall clock's progress, and in the delivery of the media stream.



Wall Clock Timeline		67	68	69	70	71	72	73	74	75
Timebase Timeline	rate 0.0	50	50	50	50	50	50	50	50	50
	rate 1.0	46	47	48	49	50	51	52	53	54
	rate -1.0	54	53	52	51	50	49	48	47	46
	rate 2.0	42	44	46	48	50	52	54	56	58

Figure 0.1: Basic model of time-controlled playback

Extending from this basic model, figure 0.2 shows how the playback of two independent media streams can be controlled on two independent media players in a coordinated fashion. To adapt the playback of the two media streams - for example to make the audio from one suitable for consumption with the video from the other - both the wall clocks and the media stream playback need to be coordinated between the two players. In the context of the present document, this happens by exchanging information between the two players across a home network. All mechanisms and solutions defined in the present document will build on and extend from this basic conceptual model. These solutions are not limited to audio-visual content but cover any type of timed content, for example subtitles, trigger events received in the broadcast and timed content generated locally by an application running on either of the devices (e.g. in the quiz show use case mentioned above).

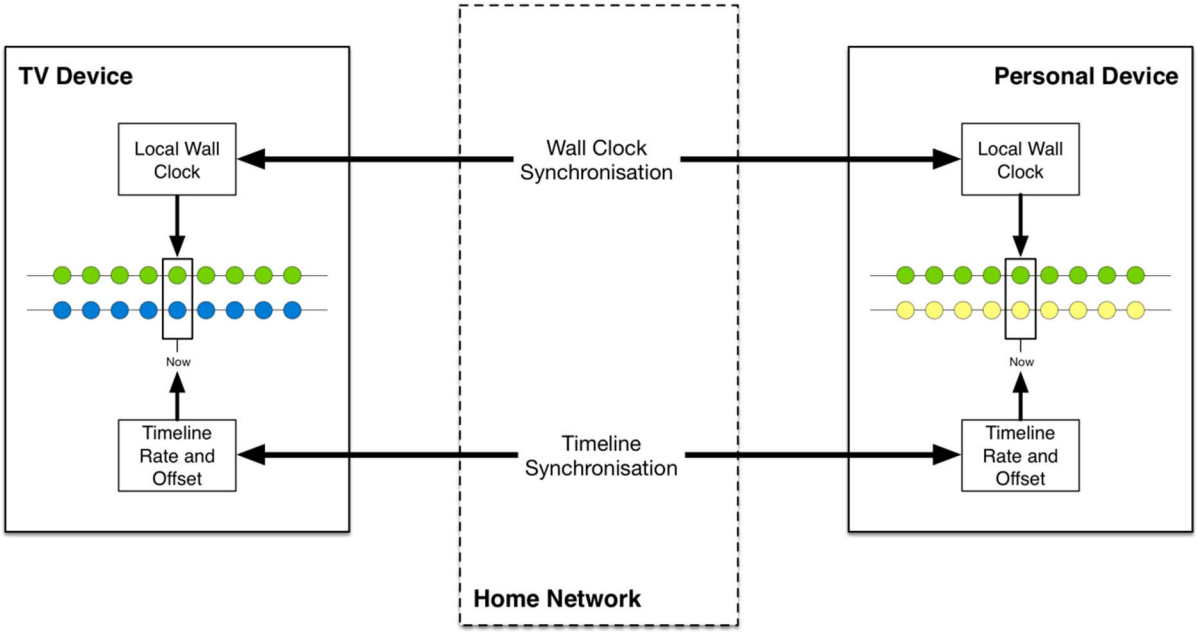


Figure 0.2: Basic model of synchronizing playback between devices

1 Scope

The present document specifies the architecture and protocols for content identification, Timeline Synchronization and Trigger Events for companion screens and streams.

The present document is applicable to:

- the interfaces between the TV Device and the Companion Screen Application:
  - interface for metadata exchange, including content identification;
  - interface for Wall Clock synchronization;
  - interface for Timeline Synchronization;
  - interface for Trigger Events;
- the interface between the Companion Screen Application and the Material Resolution Service (MRS).

2 References

2.1 Normative references

References are either specific (identified by date of publication and/or edition number or version number) or non-specific. For specific references, only the cited version applies. For non-specific references, the latest version of the referenced document (including any amendments) applies.

Referenced documents which are not found to be publicly available in the expected location might be found at http://docbox.etsi.org/Reference.

NOTE: While any hyperlinks included in this clause were valid at the time of publication, ETSI cannot guarantee their long term validity.

The following referenced documents are necessary for the application of the present document.

[1] IETF RFC 3986 (2005): "Uniform Resource Identifier (URI): Generic Syntax".

NOTE: Available at http://www.ietf.org/rfc/rfc3986.txt.

[2] ETSI TS 102 323 (V1.5.1): "Digital Video Broadcasting (DVB); Carriage and signalling of TV-Anytime information in DVB transport streams".

NOTE: Available at http://www.etsi.org/deliver/etsi_ts/102300_102399/102323/01.05.01_60/.

[3] ETSI TS 102 851 (V1.3.1): "Digital Video Broadcasting (DVB); Uniform Resource Identifiers (URI) for DVB Systems".

NOTE: Available at http://www.etsi.org/deliver/etsi_ts/102800_102899/102851/01.03.01_60/.

[4] IETF RFC 5234 (2008): "Augmented BNF for Syntax Specifications: ABNF".

NOTE: Available at http://www.ietf.org/rfc/rfc5234.txt.

[5] ETSI TS 102 034: "Digital Video Broadcasting (DVB); Transport of MPEG-2 TS Based DVB Services over IP Based Networks".

NOTE: Available at http://www.etsi.org/deliver/etsi_ts/102000_102099/102034/.

[6] ISO/IEC 13818-1:2015: "Information technology -- Generic coding of moving pictures and associated audio information -- Part 1: Systems".