HAL
open science

Hybrid Broadcast Services using MPEG DASH

J. Le Feuvre, Cyril Concolato

To cite this version:

J. Le Feuvre, Cyril Concolato. Hybrid Broadcast Services using MPEG DASH. MediaSync, Oct 2013, Nantes, France. hal-02287386

HAL Id: hal-02287386

https://telecom-paris.hal.science/hal-02287386v1

Submitted on 5 Oct 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Hybrid Broadcast Services using MPEG DASH

Jean Le Feuvre, Cyril Concolato

Telecom ParisTech, Paris, France

E-mail: {jean.lefeuvre, cyril.concolato}@telecom-paristech.fr

Abstract: **The emergence of new multimedia services such as super high definition or free viewpoint TV will put a very large bandwidth pressure on the broadcast networks. In order to offer these services while maintaining quality, service provider will have to use other delivery mechanisms. In this context, enhancing broadcast services with broadband media may be one key to low cost deployment of new services. This paper studies the problem of hybrid broadcast and broadband delivery of audio-visual content and reviews existing solutions. With the growing importance of HTTP streaming services, and the adoption of MPEG-DASH by the TV industry, a solution for hybrid broadcasting based on this standard is proposed. Implementation is presented and evaluated, and future work directions are discussed.**

Keywords: Hybrid, broadcast, broadband, adaptive streaming, Future TV

1 INTRODUCTION

With the emergence of new multimedia technologies such as ultra-HD (UHD) or 3D displays, the TV broadcasting industry is facing new challenges regarding the optimization of the broadcast network, be it satellite or terrestrial. Moreover, user demand for richer services, such as free viewpoint selection or high definition audio enhancement (language selection, multichannel streams), also impacts the network occupancy. Although new broadcasting standards such as DVB-S2 [1] or DVB-T2 [2] offer more bandwidth, and the new HEVC [3] video compression standard allows for higher compression ratios, long is the road until broadcaster may offer full UHD 3D services to their customers.

Another interesting aspect lies in the non-uniformity of the end-user equipment. 3DTV are not widespread; UHD displays are at the experimental stage and will not hit the mass consumer market any time soon. Moreover, not all users equipped with 3D TVs are willing to pay for 3D broadcast services, limiting their 3D experience to home video. Considering all these issues, one might easily understand the reluctance of TV broadcasters to waste precious broadcast bandwidth only to address a very small subset of their users.

Nevertheless, if this subset is willing to pay for such services, a solution should be found to help for the deployment of new services. With the growing adoption of connected TVs or set-top boxes, it is worth investigating whether enhancement to broadcast services could be delivered through a broadband link; this would allow broadcasters to address their complete user base through their broadcast network, thereby minimizing their broadcast costs, while offering premium services of various kinds to specific users; we can cite many use cases:

- Switching a SD/HD service into a HD/UHD using SVC or the upcoming HEVC Scalable standard;
- Switching a 2D service into a 3D one using MVC or service-compatible AVC enhancement stream;
- Switching views in a live sport event;
- Switching audio to any possible language or high definition version.

The common point between these scenarios is the ability to synchronize the broadband and broadcast streams in a very fine way, requiring frame-accurate synchronization for 3D or spatial enhancements, or even at the audio sample level if using a multistream audio representation.

Delivering a multimedia content over both broadcast and broadband is usually referred to as “hybrid broadcasting”, and the synchronization process needed is referred to as “hybrid synchronization”. The work presented in this paper is a follow-up of previous work [4], and presents a solution for such hybrid use cases based on the new MPEG DASH standard.

This paper is organized as follows; in Section 2, we will review the specificities of hybrid broadcasting and the problems induced by the broadcast environment; in Section 3, existing approaches to solve these issues are explained; in Section 4, a quick overview of MPEG DASH in the context of hybrid networking is presented; our solution is explained in section 5; Section 6 discusses the experiments and Section 7 concludes this paper and presents future work directions.

2 HYBRID BROADCASTING

As shown in [4], hybrid broadcasting raises two interesting questions: the bootstrapping resolution, or how a client discovers that media distributed over different deliveries can be consumed together; and the synchronization of these media.

Corresponding author: Jean Le Feuvre, Telecom ParisTech, 46 rue Barrault, 75013 Paris, France, jean.lefeuvre@telecom-paristech.fr

2.1 Discovery of Hybrid Enhancement

There are two general situations for the discovery of hybrid content: program guide indications or inband indications. In the first case, add-on enhancements are available through broadband information, such as an electronic program guide or any proprietary mechanism in the DTV receiver, in order to describe which broadcast channel may be enhanced. In the second case, signalling is included in the broadcast stream, for example through an HbbTV [7] application, to provide the description of the broadband media.

In order not to rely on service guide or third-party servers, we based our solution on inserting the enhancement description in the broadcast content in the form of an URL to minimize bandwidth occupancy.

2.2 Media Synchronization

Multimedia synchronization is a fairly simple process that can be described as follows: media frames (video, audio, subtitles, …) are compressed and assigned a timestamp, indicating the time at which the unit has to be presented on the display or at the output of the speakers. These timestamps refer to a timeline, which is usually not the intrinsic media timeline (e.g. time 0 for the first frame of a movie), but rather expressed with a random origin and frequency. Some standards such as RTP use one timeline per media, and indicate the relationship between timelines through a specific mechanism (RTCP reports for the RTP case). In the broadcast world, MPEG-2 TS is used as the media unit delivery mechanism, and its timing mechanism can be described as follows:

- All media streams of a program (TV channel) refer to the same timeline;
- This timeline is described by the Program Clock Reference, and is computed through PCR timestamps inserted in the transport stream, usually carried by the video stream;
- Time origin of the PCR is meaningless;
- PCR may suffer discontinuities at any time.

The main issue with MPEG-2 TS in hybrid scenarios comes from this last property; most of the deployed equipment (statistical multiplexer, transcoders, splitters…) will rewrite PCR and timestamps throughout the chain, without the content provider being aware of this transformation. In order to achieve hybrid synchronization at the frame level as previously explained, one would need to adjust the timestamp in the broadband stream to align them with the PCR. This requires real-time feedback from the equipment, for each discontinuity happening on the PCR. These modifications would then require a rewriting of the broadband content, making offline production or content distribution optimizations almost impossible. If several devices may modify the PCR in the broadcast chain, as is usually the case, such a solution is simply impractical.

Hybrid broadcasting also requires handling of broadband network jitter and downlink transmission time [4][6][9]. This is not in the scope of this paper, since we assume that the broadband enhancement is available ahead of time; our experiments were only conducted on pre-produced content. This implies that the client can fetch broadband data starting from some point ahead in time, typically several times the downlink transmission time.

3 RELATED WORK

Several approaches exist towards synchronization of broadcast and broadband content. This section reviews a few of them.

3.1 Timeline Estimation Techniques

Multiple solutions have been found to try to overcome the existing lack of synchronization support when designing add-ons to TV, without modifying existing broadcast streams.

One solution is fingerprinting [8][9]. Fingerprinting means extracting relevant parameters (the fingerprint) from an audio stream, captured by the microphone of a smartphone or tablet standing in front of an active TV set. The fingerprint is sent to a server, where it is compared with the fingerprints of all TV channels. The fingerprint match allows for:

- Identification of the channel the TV set is tuned into
- Timing information recovery, which depends on the fingerprinting parameters but can reach a precision of 10ms [8].

Fingerprinting works well to get a single timing indication, and to some extend may be used to provide near lip-sync support. However, it is not precise enough to cover frame-accurate precision, needed in hybrid delivery of hierarchically coded scalable streams, for which one frame delay between streams will lead to decoding errors. It is also sensitive to ambient noise, and server load may delay responses beyond usability with a growing number of users.

3.2 Best Effort Timeline Techniques

Another approach consists in using existing timing information from the broadcast content. As explained previously, using the PCR values is not possible due to clock discontinuities. [4] shows that it is possible to use UTC timing conveyed in DVB networks through the TDT table. This information is sufficient for some use cases, such as triggering recordings or overlaying some applicative data (stock exchange, interactive quiz…) with loose synchronization requirements. However, existing equipment do not maintain a strict synchronization between the multiplex time (therefore the PCR) and the TDT; drift can be as high as couple of seconds in existing DVB-T networks, as shown in Figure 1.

This implies that frame-accurate synchronization cannot be achieved this way without impacting most of the broadcast chain.

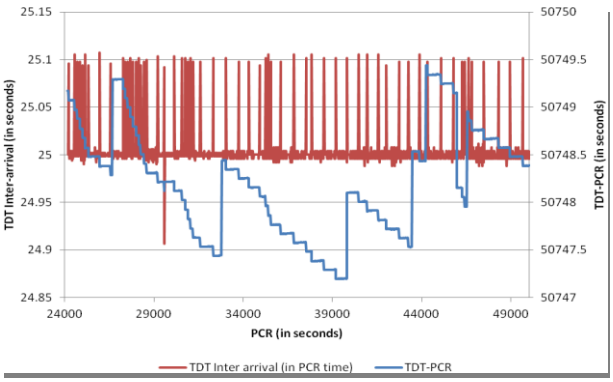

Figure 1: TDT/PCR drift in DVB-T multiplex

Moreover, this solution would require very fine synchronization between the broadcast chain and the broadband provider to avoid drift in their respective UTC clock. Finally, using a UTC based timeline makes synchronization more complex, since broadband servers have to signal how to map the broadcast UTC time to the timeline of pre-produced broadband content.

3.3 Timeline Data Streaming

Another approach to solve the synchronization issue is to send another timeline in the MPEG-2 TS, along with the PCR. Such a timeline is not modified by PCR discontinuities and can therefore be used as a synchronization reference by the broadband content. This approach is standardized in [11], and covers many use cases including recomposing of a movie timeline after removal of ads. The design principle is simple: the timeline is sent as a PES stream, and the payload of the PES packet contains the timeline value corresponding to the PTS value coded in the PES header. This ensures that all media (audio, video) timestamps can be transposed in this additional timeline regardless of PCR discontinuities. Since this timeline is inserted by the content producer and is not altered by the broadcasting chain, broadband content can be synchronized in the most accurate way desired (frame accuracy, audio sample accuracy) by selecting a proper resolution for this timeline. This work has been demonstrated in [9].

While this approach fits our needs, it still lacks tools to identify the location of the broadband data, as well as accurate description of the current TV show. Our solution is based on this design principle, but proposes MPEG-DASH compatible tools to enable hybrid broadcasting.

4 A DASH APPROACH

The DASH-oriented design of our solution was driven by two arguments:

- DASH is designed for delivery and synchronization of multiple media streams over various path, and has all the tools required to synchronize streams coming from different broadband servers;
- DASH is getting a growing adoption by TV standards such as HbbTV [10], and it should not be long until TVs are DASH-enabled.

While the standard is designed for HTTP based delivery, it is being adopted in other contexts such as mobile broadcasting (eMBMS), thanks to its expressivity. Especially, DASH has done a fairly complete review of how hierarchically coded streams and alternate streams can be described in an efficient way, covering both scalable coding, multiview coding and service-compatible 3D coding. Another interesting feature of DASH is its ability to handle streams coming from different origins, possibly in different formats, and therefore not using the same timeline to express their timestamps. DASH provides tools to map timelines of each media to a global time, even if these timelines do not share the same time origin. The presentation time of each media depends on the underlying media format, MPEG-2 TS or MP4. For MPEG-2 TS case, the presentation time info is usually conveyed through segment indexes, which are not present in the TS. When not present, clients rely on PCR information and period start time, which is problematic in our scenario as explained in Section 2. By overriding the notion of presentation time in live broadcast with the approach given in the previous section, we can provide a continuous TS timeline which can be mapped to other media described in the DASH session, regardless of PCR re-stamping or discontinuities happening in the network.

5 DASH HYBRID BROADCASTING

Our solution consists in delivering DASH MPD location inside the broadcast stream for each program carry the PresentationTime of the program as well as the period information in a dedicated PID. An overview of the architecture of this hybrid system is given in Figure 2.

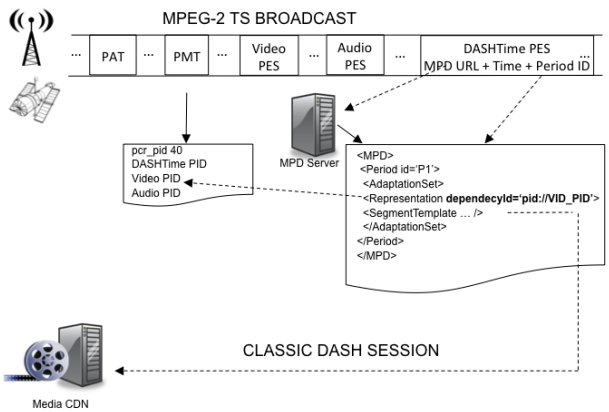

Figure 2: Proposed DASH-based hybrid broadcasting architecture

In a first design, the DASH information was embedded in an MPEG-2 TS section, such as PMT or AIT. However, this impacts existing multiplexers, as they have to maintain tight synchronization between PES data and PMT/AIT, especially when sudden period switch happen. We therefore decide to embed the DASH description within the PES stream carrying the timeline information to limit the impact on the infrastructure and provide more robustness in case of DASH period switching. The link between the MPD media streams and the MPEG-2 TS depends on the nature of the enhancement. For non-

scalable media enhancement, no explicit link given, and media composition is decided by the digital TV terminal. For scalable media enhancement (spatial, 3D), the **dependencyId** attribute of the enhancement layers described in the MPD are set to point to the base layer in the MPEG-2 TS, through a custom URL scheme *pid://PID*. The solution has therefore no impact on the XML validation of the MPD and does not modify the URL resolver algorithm of the DASH client.

Table 1: DASHTime PES payload format

```
DASHTime_PES_Payload {
 bit(64) presentation_time;
 bit(64) activation_countdown;
 byte mimeType;
 byte URL_location;
}
```

We also defined a new stream type for our DASH timeline stream, called DASHTime, and defined its payload as shown in Table 1. This fixed payload is designed to fit in a TS packet, including PES header and optional adaptation fields of the TS packet. It carries the presentation time of all streams in the program, by associating this value with the PTS value given in the PES header of the packet. The DASH packet may be sent at any time, but our implementation only sends them at the beginning of each frame or after every packet containing a PCR. The packet also indicates the location of the DASH MPD as well as the current period id in order to properly handle period changes. Finally, the packet includes an activation countdown announcing that the given URL will not be valid until the indicated time, allowing a client to prefetch media data before the actual period switch occurs. Once the broadcasted content can be identified in the MPD, the client can process the MPD as usual, following DASH rules for representation selection and MPD updates, and a regular DASH session can, then take place for broadband media, as shown in Figure 2.

Our solution provides a generic mechanism allowing broadband add-ons to be synchronized with MPEG-2 TS broadcast, regardless of the add-on type; MPEG DASH is of course a main use case for our solution, but other enhancements such as plain files, RTP streams or MMT content could be synchronized with this proposal.

6 EXPERIMENTS

We have experimented with this design using multicast TS as a broadcast simulator and an HTTP server, using [12] as the underlying platform for content generation (DASH and TS) and playback. Our experiment shows that the timeline reconstruction works perfectly and gives us a frame accurate playback. However, we only experienced with non-live broadband content, and the player always had the ability to fetch broadband data ahead of time, typically one second. This increases the hybrid setup delay, but ensures that broadband data to add was already available for decoding before broadcast data arrived. As in most broadcast use cases, we did not want to consider the pausing of the broadcast stream to cope with initial buffering. Future work will investigate the possibilities to perform this experiment in live conditions, with the smaller production delay possible.

Another limitation of our system is the live-only design; we still have to investigate how performing trick-modes on recorded or time-shifted broadcast could be enhanced with the tools that DASH provides.

7 CONCLUSION

In this paper, we have presented some use cases for hybrid broadcasting, reviewed the different approaches towards solving these scenarios and explained how the MPEG-DASH standard provides most of the tools needed to fulfil them. We have proposed a light solution compatible with existing infrastructure ensuring a frame-accurate synchronization between broadcast and broadband stream. This work is under investigation for standardization by the MPEG committee, with a work plan targeting both explicit technical tools for the MPEG-2 transport stream and generic transport guidelines for achieving multi-container synchronization. In future work, we will investigate usage of DASH trick modes for broadcasted content as well as low delay live hybrid broadcasting.

Acknowledgment

This work was partially funded by the European Commission in the course of the CELTIC project H2B2VS and by the French Research Agency in the course of the OptiSaT2 project.

References

[1] A. Morello,"DVB-S2: The second generation standard for satellite broadband services," Proc. IEEE, vol. 94, no. 1, pp. 210-227, Jan. 2006.
[2] L. Vangelista, N. Benvenuto, S. Tomasin, C. Kokes, J. Stott, A. Filippi, M. Vlot, V. Mignone, and A. Morello, "Key technologies for next-generation terrestrial digital television standard DVB-T2," IEEE Communications Magazine, vol. 47, no. 10, pp. 146-153, 2009.
[3] G. J. Sullivan and J.-R. Ohm, "Recent developments in standardization of high efficiency video coding (HEVC)," Proc. SPIE, vol. 7798, Aug. 2010.
[4] C. Concolato, S. Thomas, R. Bouqueau, and J. Le Feuvre, "Synchronized delivery of multimedia content over uncoordinated broadcast broadband networks," in Proceedings of the 3rd Multimedia Systems Conference on - MMSys '12, 2012, p. 227.
[5] Sodagar, I., The MPEG-DASH Standard for Multimedia Streaming Over the Internet, IEEE Multimedia, vol. 18, no. 4, pp. 62-67, Oct.-Dec. 2011
[6] U. Rauschenbach, W. Putz, P. Wolf, R. Mies and G. Stoll, "A scalable interactive TV service supporting synchronised delivery over broadcast and broadband networks", IBC Conference, September 2004.
[7] ETSI TS 102 796: "Hybrid Broadcast Broadband TV", version 1.5.
[8] J. Haitsma, T. Kalker, and J. Oostveen, "Robust Audio Hashing for Content Identification Robust Audio Hash Definition," in International Workshop on ContentBased Multimedia Indexing, 2001, vol. 4, no. 82, pp. 117-124.
[9] C. Howson, E. Gautier, P. Gilberton, A. Laurent, and Y. Legallais, "Second screen TV synchronization," in 2011 IEEE International Conference on Consumer Electronics -Berlin (ICCE-Berlin), 2011, pp. 361-365.
[10] HBBTV, ETSI Technical Specification 102 796, "Hybrid Broadcast Broadband TV"
[11] ETSI TS 102 823 "carriage of synchronized auxiliary data in DVB transport streams"
[12] Le Feuvre, J., Concolato, C., and Moissinac, J. 2007. GPAC: open source multimedia framework. In Proceedings of the 15th international Conference on Multimedia (Augsburg, Germany, September 25 - 29, 2007). MULTIMEDIA '07.