



ATSC

ADVANCED TELEVISION
SYSTEMS COMMITTEE

**ATSC Standard:
A/331:2022-11 Amendment No. 2,
“Timing/Buffer Model – IMSC1”**

Doc. A/331:2022-11 Amend. No. 2
16 February 2023

Advanced Television Systems Committee
1300 I Street, N.W., Suite 400E
Washington, D.C. 20005
202-872-9160

The Advanced Television Systems Committee, Inc. is an international, non-profit organization developing voluntary standards and recommended practices for broadcast television and multimedia data distribution. ATSC member organizations represent the broadcast, professional equipment, motion picture, consumer electronics, computer, cable, satellite, and semiconductor industries. ATSC also develops implementation strategies and supports educational activities on ATSC standards. ATSC was formed in 1983 by the member organizations of the Joint Committee on Inter-society Coordination (JCIC): the Consumer Technology Association (CTA), the Institute of Electrical and Electronics Engineers (IEEE), the National Association of Broadcasters (NAB), the Internet & Television Association (NCTA), and the Society of Motion Picture and Television Engineers (SMPTE). For more information visit www.atsc.org.

Note: The user's attention is called to the possibility that compliance with this standard may require use of an invention covered by patent rights. By publication of this standard, no position is taken with respect to the validity of this claim or of any patent rights in connection therewith. One or more patent holders have, however, filed a statement regarding the terms on which such patent holder(s) may be willing to grant a license under these rights to individuals or entities desiring to obtain such a license. Details may be obtained from the ATSC Secretary and the patent holder.

Implementers with feedback, comments, or potential bug reports relating to this document may contact ATSC at <https://www.atsc.org/feedback/>.

Revision History

Version	Date
Amendment approved	16 February 2023

ATSC Standard: A/331:2022-11 Amendment No. 2, “Timing/Buffer Model – IMSC1”

1. OVERVIEW

1.1 Definition

An Amendment is generated to document an enhancement, an addition, or a deletion of functionality to previously agreed technical provisions in an existing ATSC document. Amendments shall be published as attachments to the original ATSC document. Distribution by ATSC of existing documents shall include any approved Amendments.

1.2 Scope

This amendment is in response to [New Project Proposal N-035](#). It addresses the need for special timing and buffering consideration for IMSC1 tracks, adds required timing constraints, and introduces possible implementations.

1.3 Rationale for Changes

Unlike video and audio, IMSC1 cannot be encoded until an entire Segment, e.g., 2s, has been ingested by an encoder. Not recognizing this has allowed encoders/packagegers to output IMSC1 that results in the Receiver being 2 seconds behind the video and audio for IMSC1 decoding. Either that or the Receiver taking an extra 2 seconds for service acquisition due to Receiver buffering. This amendment provides a recommendation for how to correct this.

1.4 Compatibility Considerations

The changes described in this document are only recommendations and there is no evidence that there would be any impact on interoperability.

2. LIST OF CHANGES

Change instructions are given below in *italics*. Unless otherwise noted, inserted text, tables, and drawings are shown in **blue**; deletions of existing text are shown in ~~red-strikeout~~. The text “[ref]” indicates that a cross reference to a cited referenced document should be inserted.

2.1 Change Instructions

Add new Section 8.1.1.7

8.1.1.7 Timing and Buffer Model Considerations with IMSC1

Notwithstanding the details of sections 8.1.1.1 ~ 8.1.1.6, there are some additional considerations when transmitting IMSC1. Unlike video and audio tracks, IMSC1 text tracks require that the Receiver acquire all LCT packets of the Segment before decoding can begin. The IMSC1 payload consists of a single IMSC1 “document” where all structures must be available to the Receiver. This is different than video and audio where decoding (and presenting) can start after a modest amount of data has been acquired, but considerably less than a typical 2-second duration Segment.

Nominal coding of video, audio and IMSC1 without special handling of IMSC1 LCT packets is shown in Figure [X].

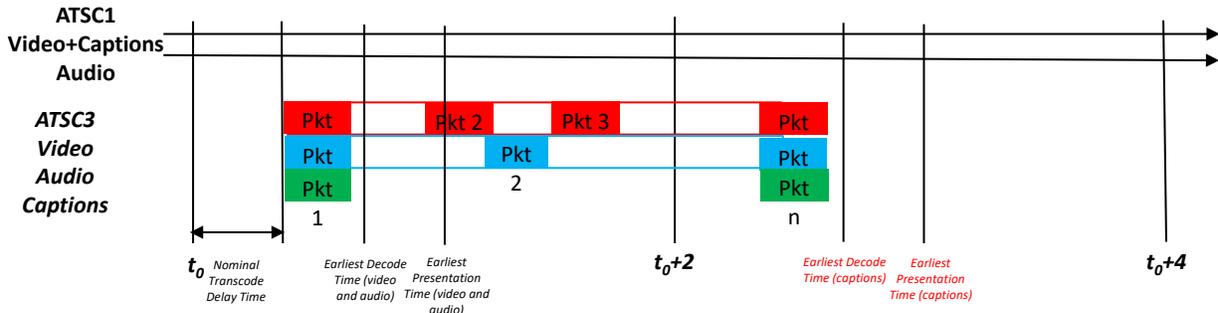


Figure X Typical Segment LCT packet emission

The last LCT packet of an IMSC1 Segment cannot be transmitted until after receipt of the last of the (2-sec) input captions. Therefore, IMSC1 decoding and presentation times are either a Segment duration late (resulting in a 2 sec delay in caption text on screen) or the presentation time is “too late” relative to video and audio and nothing is displayed by the Receiver.

It is possible for the Receiver to compensate for this skew by buffering of full segments and delaying the decoding and presentation of video and audio until after the last IMSC1 LCT packet. However, this results in an unsatisfactory viewer experience of poor service acquisition time, delaying it by a Segment duration, e.g. 2 more seconds after acquiring the beginning of a Segment (1 second on average for a 2-second Segment).

To address this in the encoding/packaging prior to emission, in addition to other transmission requirements (notably buffer model constraints):

- 1) All Segment boundaries for all tracks that are synchronized (nominally video, audio, and captions) shall be aligned in time (start, end or both) with the track with the longest duration Segment;
- 2) All LCT packets shall be contained within their Segment start and end times; and
- 3) For each Segment with IMSC1 caption data, the last LCT packet carrying the last data of the IMSC1 document shall be emitted in the first 400ms of the Segment.

One encoder/packager side solution is that the encoder/packager could delay the output of video and audio to be aligned with one full Segment of IMSC1 coded as shown in Figure [Y].

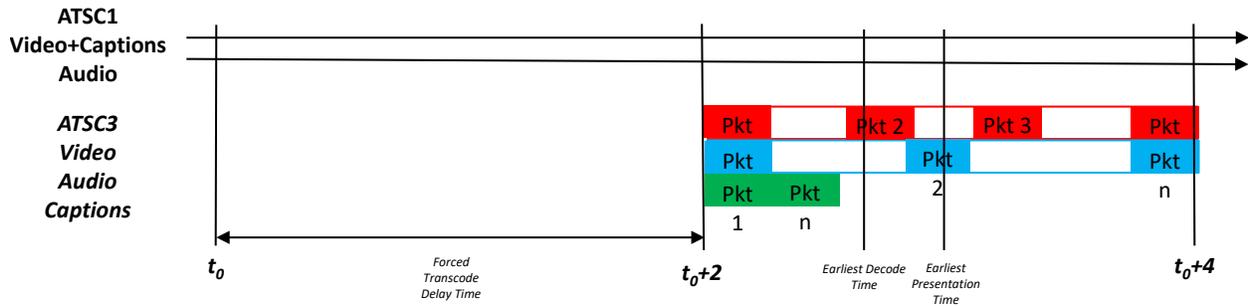


Figure Y Recommended Segment LCT packet emission

Note that the IMSC1 Segment is transmitted up front in the Segment timeline and can be decoded and presented at the same time as video and audio. From the Receiver (and viewer) perspective the service can be acquired rapidly and IMSC1 remains in sync with video and audio. This might increase the encoding time of Segments depending on current encoding delays today, and thus adversely impact low latency applications.

An alternative encoder solution is reducing the IMSC1 Segment duration to roughly the video and audio decode and presentation delay, e.g. ~400ms. However, this introduces:

- 1) notable (5x) the Receiver computation to decode and display IMSC1 compared to 2s Segments;
- 2) does not conform to CMAF minimum Segment duration of 960ms (but could be corrected by using CMAF Fragments); and
- 3) introduces VBR coding for IMSC1

IMSC1 decoding is not necessarily performed with hardware assistance and might be performed with a general purpose processor with limited performance. In order to simplify processing and allow adequate time for decoding while minimizing the necessity of receiver-side buffering of audio and video, each IMSC1 document (sample) shall be delivered in its entirety at least 1 second before its associated presentation time.

– End of Document –