

Seamless Sequential Playback (Best Practice)

CONTENTS

1	Seamless Sequential Playback Model	4
1.1	Summary.....	4
1.2	References.....	4
1.3	Definitions	4
2	Track Constraints.....	7
3	Player Behavior.....	8
Annex A	Examples.....	9



This work is licensed under a [Creative Commons Attribution 3.0 Unported License](https://creativecommons.org/licenses/by/3.0/).

NOTE: No effort is being made by the Motion Picture Laboratories to in any way obligate any market participant to adhere to this specification. Whether to adopt this specification in whole or in part is left entirely to the individual discretion of individual market participants, using their own independent business judgment. Moreover, Motion Picture Laboratories disclaims any warranty or representation as to the suitability of this specification for any purpose, and any liability for any damages or other harm you may incur as a result of subscribing to this specification.

REVISION HISTORY

Version	Date	Description
1.0	June 12, 2015	Initial publication

1 SEAMLESS SEQUENTIAL PLAYBACK MODEL

This document describes the seamless playback model for playing Playable Sequences with Clips designed for seamless playback. See *Common Metadata Media Manifest [Manifest]*, Section 5.2.

Seamless sequential playback allows two or more Presentations or sections of a Presentation (Clip) to play in sequence without interruption.

Although some models require the simultaneous decoding of two audio tracks to ensure seamless playback, this model does not.

If properly implemented, this model will maintain A/V sync to within the duration of an audio frame over the course of the entire seamless playback portion of a Playable Sequence. Audio will also never precede video, a phenomenon that is disconcerting to a viewer.

1.1 Summary

The approach described here allows Clips within a Playable Sequence to be played seamlessly with audio never playing before the associated video and never more than one audio frame late.

This is accomplished by dropping an audio frame at the end of a clip if it would cause audio to play one frame late or later. This approach depends on properly authored clips.

1.2 References

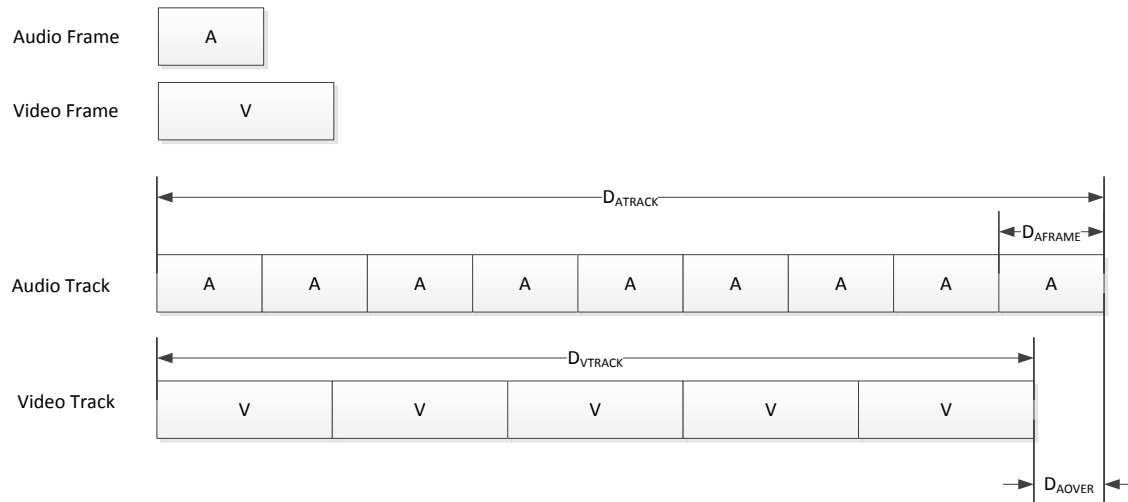
[Manifest]	MovieLabs Common Media Manifest Metadata v1.1, www.movielabs.com/md/manifest
[DMedia]	DECE Common File Format and Media Specification, www.uvcentral.com

1.3 Definitions

The terms *Playable Sequence*, *Presentation* and *Clip* are defined in Media Manifest, found at www.movielabs.com/md/manifest.

This paper concerns itself with Clips within a Playable Sequence that are flagged for seamless playback with the Clip/@seamless attribute set to 'true'.

In this model, the following diagram shows nomenclature related to a Clip.

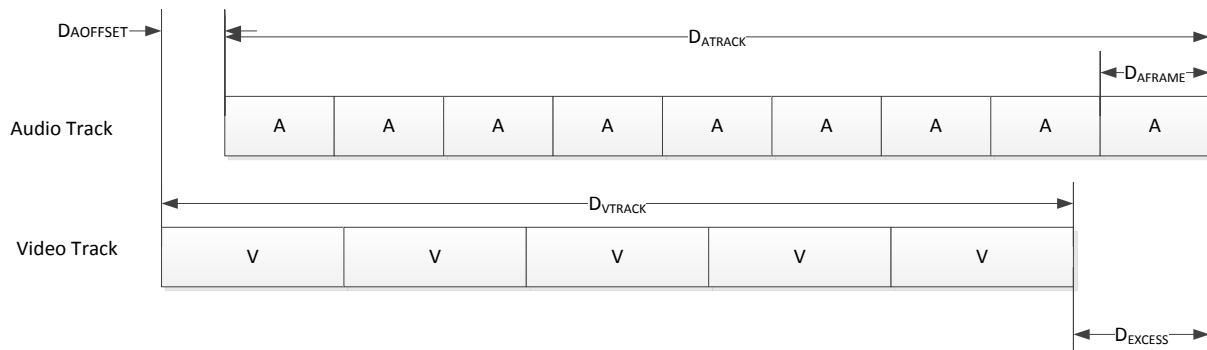


The following parameters are used

- D_{ATRACK} is the duration of the entire audio track
- D_{VTRACK} is the duration of the entire video track
- D_{AFRAME} is the duration of each audio frame in the audio track. Assumed to be constant.
- $D_{AOFFSET}$ is the initial offset of audio in a Clip.
- $D_{AOVER} = D_{ATRACK} - D_{VTRACK}$, the duration that audio exceeds video. Since the audio track must be at least as long as video, this is always non-negative.

Parameters are expressed as units of time, although the exact unit is not relevant.

If a Clip is taken from the middle of a Presentation audio might be offset from the video as follows:



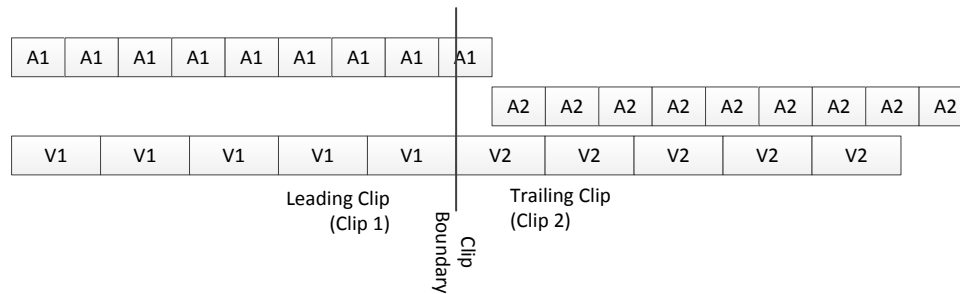
In this case the following parameters become evident:

- $D_{AOFFSET}$ is the initial offset of audio in a Clip. If Clips start simultaneously with video, $D_{AOFFSET} = 0$.

- D_{EXCESS} is the time that audio would exceed video in time if the entire D_{ATRACK} was played out. If audio and video in a clip start simultaneously, $D_{EXCESS} = D_{AOVER}$.

Note also that D_{AOVER} (i.e., $D_{ATRACK} - D_{VTRACK}$) is also equal to $D_{EXCESS} - D_{AOFFSET}$.

The following diagram defines nomenclature related to playing Clips in sequence. The following shows two Clips (1 and 2) playing in sequence seamlessly. The *Clip Boundary* is the point at which the one Clip's video ends and the next Clip's video begins.

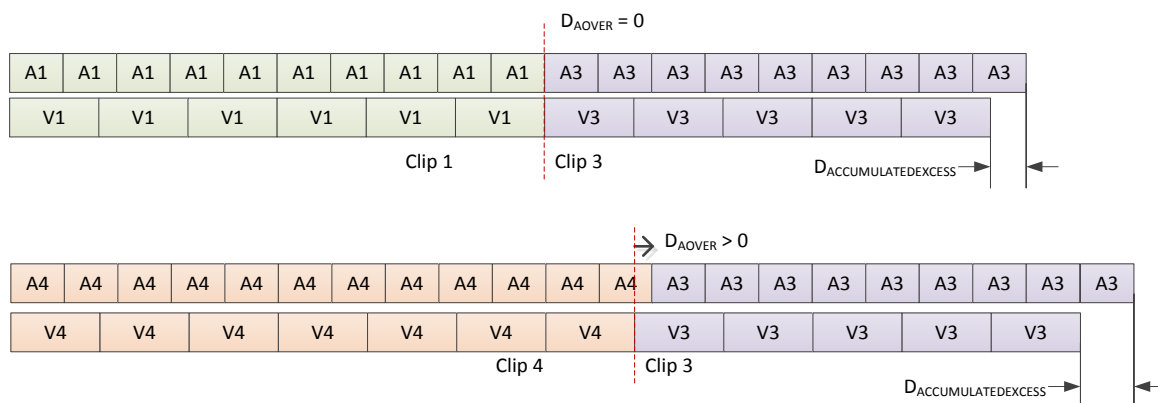


The first of successive Clips in a seamless portion of a Playable Sequence is called the *Leading Clip*. The second of successive Clips in a seamless portion of a Playable Sequence is called the *Trailing Clip*. In the diagram, Clip 1 is the Leading Clip and Clip 2 is the Trailing Clip.

For this discussion, the following parameters are defined here:

- $D_{ACCUMULATEDEXCESS}$ is the amount of time that the end time of the last audio frame of the Leading Clip extends beyond the end time of the last video frame of the Leading Clip.

Consider the following illustration (note that 'Clip 1', 'Clip 2', etc. are just names for those clips):



In the above illustration, seamless playback is shown in two cases. In the first, D_{AOVER} for the clip called 'Clip 1' is 0, so $D_{ACCUMULATEDEXCESS}$ is D_{AOVER} for the clip called 'Clip 3'. In the second, both Clips have $D_{AOVER} > 0$, so $D_{ACCUMULATEDEXCESS}$ is the sum of D_{AOVER} for the two Clips.

2 TRACK CONSTRAINTS

Track constraints are given in the context of the Digital Common Container (DCC) [DMedia]. However, principles can be applied to other formats.

Audio and Video Tracks in all Clips in a seamless portion of a Playable Sequence SHALL be constructed such that D_{ATRACK} is always equal to or greater than D_{VTRACK} . That is, the audio track is never shorter than the video track. Note that this constraint is true regardless of the start time of the audio relative to the video and could cause there to be audio frames that play after the end of video.

Audio and Video Tracks in all Clips in a seamless portion of a Playable Sequence SHALL be of the same Media Profile and Delivery Target as defined in [DMedia] Annex A-C; and SHALL have all the same ecosystem-specific constraints, such as defined in [DMedia], Annex E.

All tracks in a seamless portion of a Playable Sequence SHALL use the same codec. That is, all audio tracks in all Clips in a seamless portion of a Playable Sequence SHALL use the same audio codec, and all video tracks in all Clips in a seamless portion of a Playable Sequence SHALL use the same video codec.

The audio track within a Clip SHALL start less than one audio frame duration after video start within the same clip. That is, audio and video can start simultaneously. If audio starts after video, it must start no later than one audio frame duration after video start. Audio cannot start before video.

Entry points for audio and video SHALL be the beginning of a fragment.

Note that Timed Text tracks are tied to video frames so events will continue to start and end with the associated video frame.

Note that some authors have found it beneficial to fade audio to silence at the end of each clip to avoid possible audible artifacts. Authors should select clip boundaries with discretion.

3 PLAYER BEHAVIOR

The player behavior is as follows.

At the start of playback, the player **SHOULD** synchronize audio and video. That is, if audio follows video ($D_{\text{OFFSET}} > 0$), playback **SHOULD** maintain this offset. However, the player **MAY** start audio and video simultaneously. Note that although the playback algorithm maintains audio-video synchronization be less than or equal to the duration of an audio frame ($D_{\text{ACCUMULATED EXCESS}} \leq D_{\text{FRAME}}$) in situations of overlapped timeframes, better synchronization can be maintained by simultaneously decoding audio in the previous clip with audio of following clip.

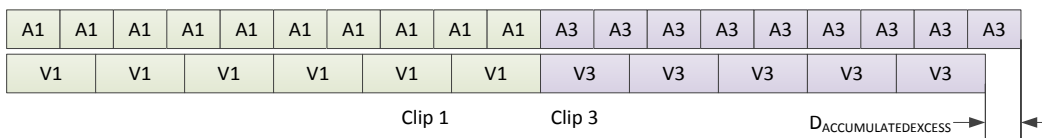
The player **SHALL** calculate $D_{\text{ACCUMULATED EXCESS}}$ in accordance with the definitions.

At the end of each Clip, the player **SHALL** discard (not play) all audio frames that exceed video playback by one frame or more. That is if $D_{\text{ACCUMULATED EXCESS}} > D_{\text{FRAME}}$, the last audio frame of the Clip is discarded.

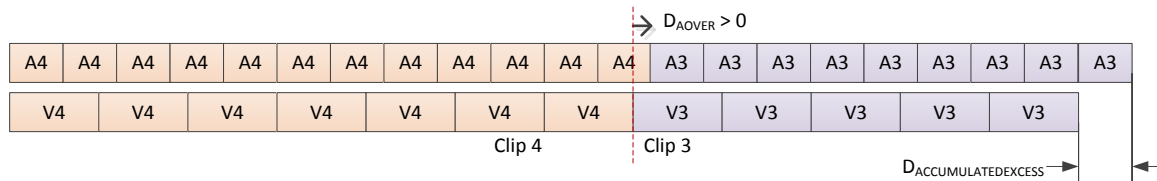
ANNEX A EXAMPLES

The following illustrates three cases of Clip boundary in a seamless portion of a Playable Sequence.

In the following case, at the end of Clip 1, audio and video tracks end at the same time so no frames need be discarded. At the end of Clip 3, $D_{\text{ACCUMULATED EXCESS}} < D_{\text{AFRAME}}$, so no frames need be discarded.



In the following case, because $D_{\text{AOVER}} (=D_{\text{ACCUMULATED EXCESS}}) < D_{\text{AFRAME}}$ no frames are discarded after Clip 4. However, at the end of Clip 3, $D_{\text{ACCUMULATED EXCESS}} \geq D_{\text{AFRAME}}$ the last audio frame of Clip 3 must be discarded before the next Clip can be played.



The following example is similar to the example above, but includes another clip:

