SCTE·ISBE CABLE-TEC EXPO 2018 | ATLANTA, GA OCTOBER 22-25

2018 Fall Technical Forum SCTE·ISBE • NCTA • CABLELABS

Efficient Session-Based Watermarking for ABR TV using Sample Variants

A Technical Paper prepared for SCTE•ISBE by

Raj Nair
VP Technology
Ericsson Media Solutions
43 Nagog Park, Acton MA 02420
978-844-5100
raj.nair@ericsson.com

Prabhu Navali
Engineering Manager
Ericsson Media Solutions
43 Nagog Park, Acton MA 02420
978-844-5100
prabhu.navali@ericsson.com

Table of Contents

List of Figures

Introduction

Session-based watermarking is on track to becoming a requirement for premium content (4K, Early Window, etc). A common approach to watermarking for ABR delivery is to use manifest files to create streams that are uniquely identified by the segment sequences. However, the storage requirements of the content is a multiple of the number of variants of a segment. In this paper, we discuss an efficient way of creating watermarked segment variants using the Sample Variants approach. Here only a small percent storage overhead is needed for each segment variant. Another advantage is that this approach scales to handle a larger population of viewers by using multiple variants with the same resource cost.

Content

1. Definitions

Sample Variant – This is an alternate representation of a sample that may be used to derive an alternate representation of a media item.

Segment Variant – This is an alternate representation of a segment that is used to derive an alternate representation of a media stream. A typical usage of a Segment Variant is for carrying a session watermark.

Session-based Watermarking – A type of watermarking in which every session has a unique watermark. The mark itself carries a payload that typically will not carry any PII (Personal Identifying Information); but, will most likely be traceable to a specific account or class of device streaming the content together with other relevant session data.

2. Overview

Session-based watermarking is seen as one of the way to address the problem of the "analog hole" in which the availability of high-definition displays and cameras has exacerbated the unauthorized redistribution of premium content.

Modern streaming services use ABR (Adaptive Bit Rate) technology such as HLS or MPEG DASH to stream content. In both these streaming methods, content is packaged into playable files, called segments, using a container format and streamed via http to end user devices such as Roku, AppleTV, or mobile devices. In ABR, the most common approach to implementing session watermarking is to use 2 segment variants (commonly known as A/B scheme) to create distinguishable streams using individualized manifest files resulting in the unique permutation of the variant segments used in the realization of individualized streams. The resulting overhead is 100% with no possible tradeoff for detection time vs overhead. In this paper, a more scalable approach is presented that uses sample variants for watermarking.

2.1. New scalable workflows

Recent advances in MPEG for CMAF and Sample Variants 2nd edition – allows for scalable and efficient session based watermarking workflows for delivery of on-demand and Live/Linear content for ABR TV services. The key idea behind this approach is that not all segments need to be replicated in order to arrive at a variant representation. Typically, only 10% of the segments need to have carry variants. Moreover, a tradeoff exists between overhead and detection speed. Hence, it should be possible to considerably reduce

overhead by using variants to create alternate representations corresponding to the watermarked variations.

CMAF is a newly emerging standard that uses ISOBMFF file format for media that can be used interchangeably with either an HLS-style m3u8 manifest file or MPEG DASH-style MPD. In addition, CMAF uses the 'cbcs' version of partial encryption (described below). Thus it is possible to now combine CBCS with CMAF to create a very efficient way of doing session watermarking. This is the thesis of the proposal in this paper.

Common Encryption supports 2 types of encryption modes: CBC (Cipher Block Chain) and CTR (Counter). Further, each of these modes have a partial encryption scheme called respectively, 'cbcs' and 'cens' subsample encryption scheme. In both of these modes, only a small percentage (typically 10%) of samples are encrypted.

SVNE is a new addition to the ISOBMFF file format and is an update to the existing Sample Variant extension. It introduces a way to support multiple encryption schemes simultaneously in a media file. The key idea behind SVNE is the use of samples encrypted with either cbcs or cens. The receiving device can choose its supported format and construct a playable representation on the fly.

3. Scalable WM

In the current proposal for session watermarking, we put the above constructs together to arrive at the combination of using CMAF together with Sample Variants (SV) and session Water Marking (WM), where the individual variants are used to support multiple watermarks in a CMAF container.

Overview of the workflow for session watermarking for Video on Demand (VoD) is given below.

3.1. SV WM VOD workflow

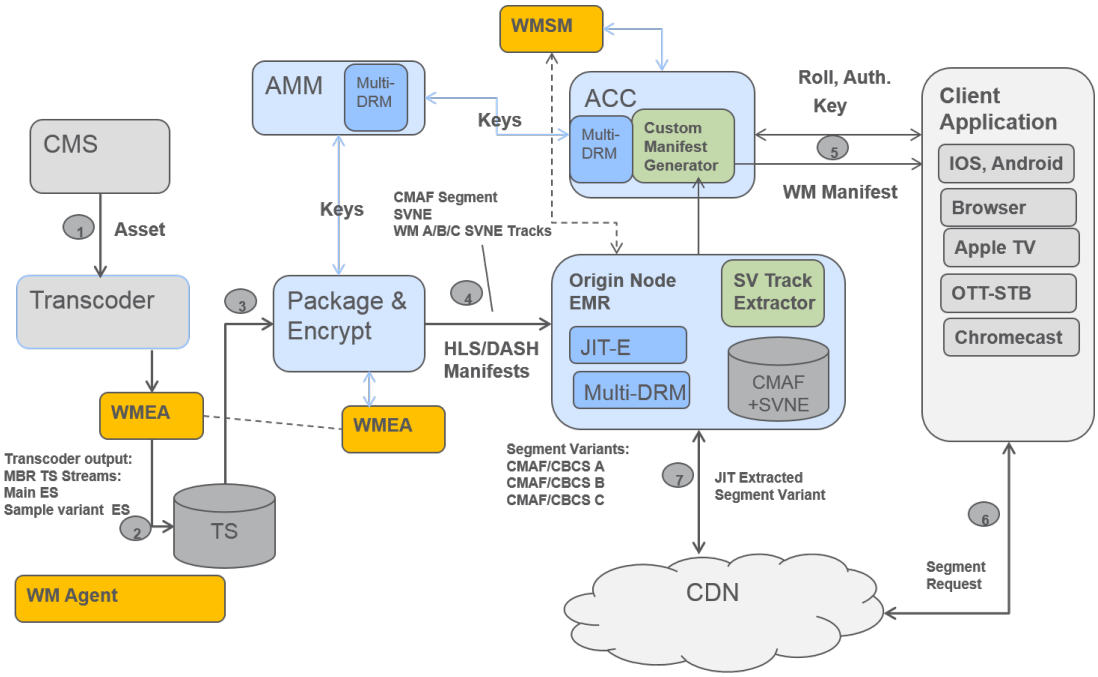

Figure 1 - SV WM VOD Workflow

In this workflow, watermarked variants of an ingested assets are created by the Watermarking Encoder Agent (WMEA) and output as a transport stream consisting of the main elementary stream together with sample variant elementary streams corresponding to the watermarks.

Each VOD asset transcoded with WMEA generates Multi-Bit Rate (MBR) files in one of two types:

1. TS A/B variant elementary streams files that can be up to K streams (k>1) or
2. TS A Stream with B TS sample variant elementary stream, where the SVNE-TS elementary streams can contain up to k sample variant elementary streams (k>1)

The next stage in the process is the packager which consults with an WMEA to produce CMAF streams with HLS and DASH manifests. The packager generates CMAF Segments containing:

1. A main track encrypted with 'cbcs'
2. SVNE WM tracks encrypted with cbcs with up to k variants from the B TS elementary stream or sample variant elementary stream as the case may be.

A Watermarking Session Manager (WMSM) produces the custom manifests for each client.

The tracks in the CMAF segment may use same or different keys for added security. In addition, the complete CMAF segment may be further encrypted with a service key (specific to each device/service provider) as a storage encryption.

Finally, CMAF Segments and manifests are uploaded to Origin Node. The Origin that plays the role of an Edge Media Router (EMR) to extract the right variants before sending to a CDN for pre-positioning content for streaming, EST, etc.

3.1.1. Client VOD workflows

There are 2 cases to consider: Client Manifest Request and Client Segment Request.

The Client Manifest Request starts with the initial authentication workflow leading to the client getting media details such as the URL to access keys as well as a custom manifest that is generated by the streaming controller using the WMSM agent. The manifest is customized by DRM, delivery format, and entitlements.

The Client Segment Request results in accessing CMAF segments from the Origin Node via the CDN cache. At the Origin itself, the EMR does a Just-in-Time Extraction (JIT-E) operation to generate the CMAF segment variant based on the DRM and delivery format.

3.1.2. Variant Segment Extraction

The variant segment extraction is done by the JIT-E function in the EMR, which may be located in the Origin. The format transformation request may be parameterized by

- WM variant A or B (or up to K)
- HLS or DASH protocol format
- CBCS or CENS encryption scheme
- Multi-DRM metadata

A segment variant is a deliverable complete segment (for playback) that is derived from the stored CMAF segment that has multiple SVNE and watermarking tracks. The EMR extracts a segment with the appropriate delivery format, encryption-scheme and DRM from the CMAF segment. There is no need to know the common encryption keys because there is no need to decrypt/re-encrypt during the JIT-E operation. HLS TS Sample AES A/B segments can be extracted from the CMAF WM A/B tracks.

Extractor is a software construct that can be embedded into a delivery node or player.

3.2. SV based Watermarking for Live ABR TV Services

The following diagram shows the workflow for SV-based watermarking for live ABR TV Services. The media preparation part of the workflow is similar except that the transcoder output is fed directly into the packager for JIT-E operation.

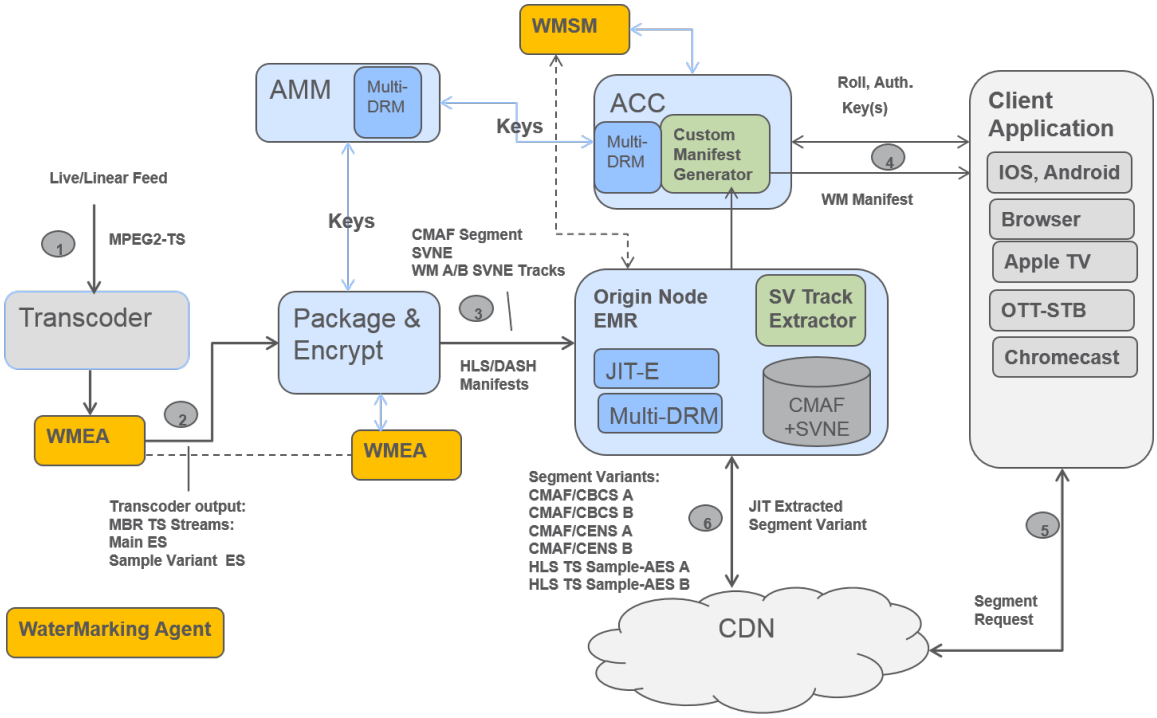

Figure 2 - SV WM Live Workflow

3.2.1. Time-shifted (nDVR) segments

In the case of time-shifted delivery, nDVR segments are generated from the stored CMAF segment with variant tracks. Delivery of the CMAF segment variant uses JIT-E.

3.2.2. Client workflows

Client work flows are similar to the VOD case with the WMSM generating custom manifests based on the DRM, delivery format and program entitlements. Likewise, client segment requests to JIT-E results in segment variant with the right delivery format, encryption and DRM.

3.3. Analysis

One drawback of the tradition approach (also known as the A/B scheme) is the need to have 2 versions of each mark-carrying segment yielding about 2^n streams where n is the number of segments carrying watermark data. However, all segments are required to have 2 variants. Hence, the overhead is 100% -- resulting in unacceptably high bandwidth and storage.

3.3.1. Tradeoff overhead vs detection

In the proposed SV WM scheme, not all segments need to have variants carrying watermarks. Hence, it is possible to derive a tradeoff based on the detection requirements needed.

Given the need to detect watermarks in a short amount of time (say 10 mins) and 2-second segments, with k=2 there is a need for at least 40 ($= 2\lceil \log_2(1{,}000{,}000) \rceil$) segments out of 300 with watermarks for a concurrent viewing audience of a million streams resulting in an overhead of 13%. For a shorter detection time or larger concurrent audiences more segments need to carry marks. The overhead can be as high as 100% for a detection period of 80 seconds. For 10M concurrent viewers and a detection period of 10 mins, the overhead can be 16%.

In the proposed method, more than 2 variants can be accommodated. The additional degree of freedom supports a higher degree of concurrency or shorter detection time without incurring large overhead. For example above, with k=3, we need only 26 (($= 2\lceil \log_3(1{,}000{,}000) \rceil$), i.e., 8.7% overhead. Using a higher value of k, say 5, we can get a detection time of 30 secs (i.e., 15 segments) but with 100% overhead. The graph below shows the relationship between the number of variants and the segments needed for detection for various levels of concurrent viewers. The number of variants do not have a large effect beyond 9 and the differences in detection times reduce with larger k as expected.

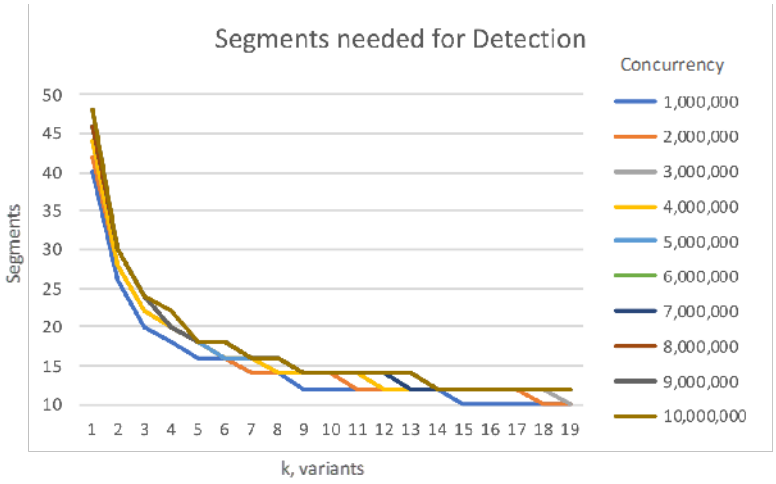

Figure 3 - Variants vs. Concurrency for Detection

Conclusion

In this paper, we have presented an improved way of doing session watermarking using sample variants. We showed details for both VoD and Live workflows for ABR TV. We also showed how this method allows more flexibility than the available A/B scheme. In particular, we demonstrated a tradeoff between

number of variants and the detection time. In other words, the overhead can be controlled for a given detection time and concurrency.

Abbreviations

ABR	adaptive bit rate
AES	Advanced Encryption Standard
CMAF	Common Media Application Format
DASH	Dynamic ABR streaming over http
DRM	digital rights management
EMR	edge media router
HLS	HTTP live streaming
ISOBMFF	ISO base media file format
JIT-E	just-in-time extraction
MBR	multi bit rate
MPD	media presentation description
nDVR	network digital video recorder
SV	sample variant
SVNE	Sample Variant Normalized Encryption
TS	transport stream
VOD	video on demand
WMEA	water marking encoding agent
WMSM	water marking session manager

Bibliography & References

ISO/IEC 23001-12 Sample Variants in the ISO base media file format, 2nd Edition

ISO/IEC 14496-12 ISO base media file format

ISO/IEC 23001-7 Common encryption in ISO base media file format files

ISO/IEC 23009-1 Dynamic adaptive streaming over HTTP (DASH), Part 1: Media presentation description and segment formats

