Media Over QUIC
Internet-Draft
Intended status: Informational
Expires: 4 June 2026

W. Law
Akamai

1 December 2025

CMSF- a CMAF compliant implementation of MOQT Streaming Format
draft-wilaw-moq-cmsf-00

Abstract

This document updates [MSF] by defining a new optional feature for the streaming format. It specifies the syntax and semantics for adding CMAF-packaged media [CMAF] to MSF.

About This Document

This note is to be removed before publishing as an RFC.

The latest revision of this draft can be found at https://wilaw.github.io/cmsf/draft-wilaw-moq-cmsf.html. Status information for this document may be found at https://datatracker.ietf.org/doc/draft-wilaw-moq-cmsf/.

Discussion of this document takes place on the Media Over QUIC Working Group mailing list (mailto:moq@ietf.org), which is archived at https://mailarchive.ietf.org/arch/browse/moq/. Subscribe at https://www.ietf.org/mailman/listinfo/moq/.

Source for this draft and an issue tracker can be found at https://github.com/wilaw/cmsf.

Status of This Memo

Table of Contents

1. Introduction

CMAF compliant MOQT Streaming Format (CMSF) is a media format designed to deliver CMAF [CMAF] and LOC [LOC] compliant media content over MOQ Transport (MOQT) [MoQTransport]. CMSF extends MSF and retains all the scope, capabilities and features of MSF including the catalog format, timeline, ABR switching and LOC support. MSF is targeted at real-time and interactive levels of live latency, as well as VOD content.

This document describes version 1 of the CMSF streaming format.

2. MSF Extension

All of the specifications, requirements, and terminology defined in [MSF] apply to implementations of this extension unless explicitly noted otherwise in this document.

3. CMAF Packaging

3.1. Initialization headers

A CMAF header is a sequence of CMAF constrained ISO BMFF boxes that do not reference any media samples, but are associated with a CMAF track and are necessary for initializing the decoding of the subsequent CMAF fragments.

The header for a given MOQT Track MUST be packaged by encoding the header using [BASE64] and then inserting that payload as the value of the Initialization data "initData" field in the catalog entry for that Track.

3.2. Switching sets and tracks

This specification defines a direct mapping between CMAF Tracks ([CMAF] Sect 3.2.1) and MOQT tracks ([MoQTransport] Sect 2.3).

CMAF switching sets are a set of one or more CMAF tracks (3.2.1), where each track is an alternative encoding of the same source content and are constrained to enable seamless track switching (3.3.9).

Each CMAF track in a switching set MUST be transmitted as a separate MOQT Track. The catalog entry for each of these tracks in the switching set MUST carry a Alternate group (altGroup) key with a common value.

The MOQT Group numbers within these switching set tracks MUST be media time-aligned. Mandating the track being media time-aligned requires that the presentation time of the first media sample contained within the first MOQT Object of each MOQT Group is identical.

3.3. Object Packaging

The payload of each Object is subject to the following requirements:

* MUST contain at least one Movie Fragment Box (moof) followed by a Media Data Box (mdat). This is equivalent to requiring that each Object hold at least one CMAF Chunk. The Media Fragment Box (moof) MUST contain a Movie Fragment Header Box (mfhd) and Track Box (trak) with a Track ID (track_ID) matching a Track Box in the initialization fragment.

* MAY contain multiple successive CMAF Chunks.

* MUST contain a single [ISOBMFF] track.

* MUST contain media content encoded in decode order.

3.4. Group Packaging

Each MOQT Group

* MUST begin with an Object containing a stream access point (SAP type 1 or 2).

* MUST contain one or more contiguous Groups of Pictures (GOPs).

* The Group boundary MUST align with a CMAF Fragment boundary. CMAF Fragments and CMAF Chunks MUST not span Groups.

3.5. Catalog description

3.5.1. CMAF packaging type

This specification extends the allowed packaging values defined in [MSF] to include one new entry, as defined in Table 1 below:

Name	Value	Reference
CMAF	cmaf	This RFC

Table 1

Every Track entry in a CMSF catalog carrying CMAF-packaged media data MUST declare a "packaging" type value of "cmaf".

3.5.2. Max SAP starting types

This specification adds two track-level catalog fields, as defined in Table 2 below:

Field	Name	Definition
Max Group SAP starting type	maxGrpSapStartingType	Section 3.5.2.1
Max Object SAP starting type	maxObjSapStartingType	Section 3.5.2.2

Table 2

3.5.2.1. Max Group SAP starting type

Location: T Required: Optional JSON Type: Number

A number indicating the maximum SAP type the MOQT Groups in the track start with.

3.5.2.2. Max Object SAP starting type

Location: T Required: Optional JSON Type: Number

A number indicating the maximum SAP type the MOQT Objects in the track start with.

3.6. Event Timelines

3.6.1. SAP Type timeline

CMSF defines a special instance of an Event Timeline track, termed the SAP Type timeline track. Its purpose is to convey information about the distribution of Stream Access Point types and their associated Earlist Presentation Times.

In the catalog, the SAP-type timeline track MUST include a ’packaging’ value of ’eventtimeline" and MUST include an ’eventType’ value of ’org.ietf.moq.cmsf.sap’.

In the SAP Type timeline JSON payload:

* The index reference MUST be ’l’ for Location

* The data field is a JSON Array containing two integers. The first integer defines SAP type with an allowed value of 0,1,2 or 3. The value 0 indicates that the Object does not start with an ISOBMFF stream access point. The value equal to 1, 2, or 3 indicates that the Object begins with a stream access point of SAP type 1, 2, or

3, respectively. When the Object is the first Object in the Group, the value MUST be equal to 1 or 2. The second integer defines the earliest media presentation timestamp, rounded to the nearest millisecond, of all media samples in the Object defined by the Location of that record.

3.6.2. SAP-type timeline track example

This shows an example of 30-fps HEVC-encoded content, in which each 4s Group beings with SAP-type 2 (i.e., the first picture in the Group is an IDR picture, while there may be one or more pictures in the Group following the IDR picture in decoding order but preceding it in output order). After 2 seconds in each Group, there is a SAP-type 3, i.e., a CRA picture, which is associated with one or more Random Access Skipped Leading (RASL) pictures. A small buffer of frames (10 frames at 30 fps) is skipped/discarded (RASL pictures) when the streaming session starts from the SAP-type 3 location. In this example, the EPT is the presentation time of the first picture after the RASL pictures in decoding order; all pictures after the RASL pictures can be fully correctly decoded and are thus presentable when the streaming session starts from the SAP-type 3 location. Note that if the streaming session starts from the start of the Group, then these RASL pictures can be fully correctly decoded and are thus presentable.

```
[
    {
        "l": [0,0],
        "data": [2,0]
    },
    {
        "l": [0,60],
        "data": [3,2100]
    },
    {
        "l": [1,0],
        "data": [2,4000]
    },
    {
        "l": [1,60],
        "data": [3,6100]
    }
]
```

4. Catalog Examples

The following section provides non-normative JSON examples of various catalogs compliant with this draft.

4.1. Simulcast video tracks - 3 alternate video qualities along with audio

This example shows catalog for a media producer capable of sending 3 time-aligned video tracks for high definition, low definition and medium definition video qualities, along with an audio track.


```
{
  "version": 1,
  "generatedAt": 1746104606044,
  "tracks":[
    {
      "name": "hd",
      "renderGroup": 1,
      "packaging": "cmaf",
      "isLive": true,
      "initData": "AAAAIGZ0eXBpc281AAA...AAAAAAAAAAAAA",
      "role": "video",
      "codec":"avc1.640028",
      "width":1920,
      "height":1080,
      "bitrate":5000000,
      "framerate":30,
      "altGroup":1
    },
    {
      "name": "md",
      "renderGroup": 1,
      "packaging": "cmaf",
      "isLive": true,
      "initData": "AAAAHGZ0eXBpc281AAA...AAAAAAAAAAAAA",
      "role": "video",
      "codec":"avc1.64001e",
      "width":720,
      "height":640,
      "bitrate":3000000,
      "framerate":30,
      "altGroup":1
    },
    {
      "name": "sd",
      "renderGroup": 1,
      "packaging": "cmaf",
      "isLive": true,
      "initData": "AAAAHGZ0eXBpc281AAA...AAAAAAAAAAAAA",
      "role": "video",
      "codec":"avc1.64000d",
      "width":192,
```

```
        "height":144,
        "bitrate":500000,
        "framerate":30,
        "altGroup":1
      },
      {
        "name": "audio",
        "renderGroup": 1,
        "packaging": "cmaf",
        "isLive": true,
        "initData": "AAAAHGZ0eXBpc281AAA...AAAAAAAAAAAAAA",
        "role": "audio",
        "codec":"mp4a.40.5",
        "samplerate":48000,
        "channelConfig":"2",
        "bitrate":67071
      }
    ]
  }
```

5. Conventions and Definitions

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

6. Security Considerations

TODO Security

7. IANA Considerations

This document has no IANA actions.

8. Normative References

[BASE64] Josefsson, S., "The Base16, Base32, and Base64 Data Encodings", RFC 4648, DOI 10.17487/RFC4648, October 2006, <https://www.rfc-editor.org/rfc/rfc4648>.

[CMAF] Standardization, I. O. for., "Information technology — Multimedia application format (MPEG-A) — Part 19: Common media application format (CMAF) for segmented media", October 2021.

[MoQTransport]
Curley, L., Pugin, K., Nandakumar, S., Vasiliev, V., and I. Swett, "Media over QUIC Transport", Work in Progress, Internet-Draft, draft-ietf-moq-transport-10, 3 March 2025, <https://datatracker.ietf.org/doc/html/draft-ietf-moq-transport-10>.

[MSF] Law, W., Curley, L., Vasiliev, V., Nandakumar, S., and K. Pugin, "WARP Streaming Format", Work in Progress, Internet-Draft, draft-ietf-moq-warp-01, 22 July 2025, <https://datatracker.ietf.org/doc/html/draft-ietf-moq-warp-01>.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/rfc/rfc2119>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/rfc/rfc8174>.

Acknowledgments

TODO acknowledge.

Author’s Address

Will Law
Akamai
Email: wilaw@akamai.com