

Media Over QUIC
Internet-Draft
Intended status: Informational
Expires: 7 May 2026

W. Law
Akamai
3 November 2025

CARP - a CMAF compliant implementation of WARP
draft-law-moq-carp-00

Abstract

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

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/carp/draft-law-moq-carp.html>. Status information for this document may be found at <https://datatracker.ietf.org/doc/draft-law-moq-carp/>.

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/carp>.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <https://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on 7 May 2026.

Copyright Notice

Copyright (c) 2025 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Revised BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Revised BSD License.

Table of Contents

1. Introduction	2
2. WARP Extension	3
3. CMAF Packaging	3
3.1. Initialization headers	3
3.2. Switching sets and tracks	3
3.3. Object Packaging	3
3.4. Group Packaging	4
3.5. Catalog description	4
3.5.1. CMAF packaging type	4
3.5.2. Max SAP starting types	4
3.6. Event Timelines	5
3.6.1. SAP Type timeline	5
3.6.2. SAP-type timeline track example	6
4. Catalog Examples	6
4.1. Simulcast video tracks - 3 alternate video qualities along with audio	7
5. Conventions and Definitions	8
6. Security Considerations	8
7. IANA Considerations	8
8. Normative References	8
Acknowledgments	9
Author's Address	9

1. Introduction

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

This document describes version 1 of the CARP streaming format.

2. WARP Extension

All of the specifications, requirements, and terminology defined in [WARP] 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:

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

CARP 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 Earliest 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.carp.sap'.

In the SAP Type timeline JSON payload:

- * The index reference MUST be '1' 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 begins 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": "cmf",
      "isLive": true,
      "initData": "AAAAIGZ0eXBpc281AAA...AAAAAAAAAAAA",
      "role": "video",
      "codec": "avc1.640028",
      "width": 1920,
      "height": 1080,
      "bitrate": 5000000,
      "framerate": 30,
      "altGroup": 1
    },
    {
      "name": "md",
      "renderGroup": 1,
      "packaging": "cmf",
      "isLive": true,
      "initData": "AAAAHGZ0eXBpc281AAA...AAAAAAAAAAAA",
      "role": "video",
      "codec": "avc1.64001e",
      "width": 720,
      "height": 640,
      "bitrate": 3000000,
      "framerate": 30,
      "altGroup": 1
    },
    {
      "name": "sd",
      "renderGroup": 1,
      "packaging": "cmf",
      "isLive": true,
      "initData": "AAAAHGZ0eXBpc281AAA...AAAAAAAAAAAA",
      "role": "video",
      "codec": "avc1.64000d",
      "width": 192,
```

```
    "height":144,
    "bitrate":500000,
    "framerate":30,
    "altGroup":1
  },
  {
    "name": "audio",
    "renderGroup": 1,
    "packaging": "cmf",
    "isLive": true,
    "initData": "AAAAHGZ0eXBpc281AAA...AAAAAAAAAAAA",
    "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>>.

[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>>.

[WARP] 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>>.

Acknowledgments

TODO acknowledge.

Author's Address

Will Law
Akamai
Email: wilaw@akamai.com