Proposal for CMSD Client-VQA keys

Dr. Urvashi Pal, William Law
Akamai Technologies
4 July, 2024

Table of Contents

---

1. Introduction

This specification is intended to enable a more accurate estimation of video quality from the viewer's perspective as compared to the traditional bitrate method. This is particularly significant given the dynamic nature of today's ad-tech landscape, with personalized content based on regions, devices, and individual profiles, which impacts the overall QoE of the viewer. Some codec companies have proprietary video quality models, therefore, this proposal defines a method for using Common Media Server Data (CMSD) [3] response headers that content preparation vendors can use to send the per-segment VQA scores to be utilized by the "Client-VQA" [1][2] model. The header processing can be performed either inside the CDN or at the client. This method also allows for detailed reporting of VQA scores for each GOP within a segment. This granularity is beneficial for LL-HLS, where each segment part corresponds to a GOP and will carry its own VQA score.

There are two primary advantages of exposing VQA (VMAF, VQM, pVMAF, PSNR, etc) scores on the client side:

a. They grant visibility into VQA for commercial and non-commercial content.
b. They facilitate matching the video quality of content between advertisements and primary content.

2. Per-Segment VQA Score Usage by CDN

The origin will publish the VQA score in a CMSD response header when the CDN requests the media segment. Once the per-segment VQA scores are received in the CDN's server logs for individual sessions, they can be extracted and compiled into an analytics dashboard by a SaaS solution (aka Client-VQA).

An alternate workflow is for the origin to still publish the VQA score in a CMSD response header, but it is ignored by the CDN and passed through to the client. The end client can then retrieve the VQA data and either send it to a 3rd party analytics solution, or pass it back to the CDN as Common Media Client Data (CMCD) fields [4].

3. Design and Naming Convention

Table 1 proposes two new CMSD [3] keys:

Table 1. CMSD new keys for per-segment VQA

Description	Key Name	Header Name	Type & Unit	Value Definition
VQA type	vqat	CMSD-Static	String or List of Strings	A string defining the type of VQA metric being reported. The allowed types are defined by the "Field name" column of Table 2. VQA types are case-sensitive. If multiple VQA metrics are being conveyed, they are represented as a List. Lists are denoted by surrounding parentheses (Unicode 0x28 and Unicode 0x29), and their values are delimited by one space (Unicode 0x20).
VQA score	vqas	CMSD-Static	Integer or List of integers	A number carrying the value of the VQA metric. The metric may be reported as an aggregate over the segment, in which case a single value exists per type. The metric may also be reported per-GOP (Group of Pictures), in which case multiple values will be reported in a list, each value applying to a GOP. If multiple types are declared, along with multiple GOPs, then the number of metrics declared per type MUST be identical

				and the position of the metrics in the list MUST match that of types to which they refer. Lists are denoted by surrounding parentheses (Unicode 0x28 and Unicode 0x29), and their values are delimited by one space (Unicode 0x20).

The processing and structure of these CMSD keys is defined by [3].

VQA scores for multiple GOPs will be listed in the ‘vqas’ value as a List in accordance with RFC8941[5]. Parsers can determine if they are extracting a single or multiple values by the presence of list parentheses. Where a video segment contains multiple GOPs, the VQA score for each GOP will be reported within the list parenthesis, for example

vqat="VMAF",vqas=(81 83)

VQA scores for multiple types also result in the ‘vqas’ value being a List, in which each list item references the ‘vqat’ type with the same ordinal list position. In this example, only a single score is reported for each type.

vqat=("VMAF" "PSNR"),vqas=(81 38)

Multiple types along with multiple GOPs may also be reported. In this case Lists are used for both the ‘vqas’ and ‘vqas’ fields. The number of GOPs referenced for each type needs to be identical and the order of the types needs to match the order of the GOP scores.

vqat=("VMAF" "PSNR"),vqas=(81 83 38 39)

3.1. VQA Types

Table 2. contains a list of all the field values that can be used with the ‘vqat’ CMSD key. Field values defined in this table are case-sensitive. All scores are normalized to be integers between [0..100||0..60] to avoid precision ambiguities.

Table 2. VQA Types Description

S.No.	Field name	Allowed values	Definition
1	VMAF	Integer [0..100]	The VMAF score for a standard profile, as defined by [6].
2	VMAFMobile	Integer [0..100]	The VMAF score for a mobile profile, as defined by [7].
3	VMAFUHD	Integer [0..100]	The VMAF score for a UHD profile, as defined by [8].
4	VMAFHD	Integer [0..100]	The VMAF score for an HD profile, as defined by [7].
5	PSNR	Integer [0..60]	The PSNR score for a standard profile, as defined by [9].
6	SSIM	Integer [0..100]	The SSIM score [10] multiplied by 100 and rounded to the nearest integer.
7	ATEME VQA Metrics	Integer [0..100]	ATEME proprietary VQA model
8	EQMnr	Integer [0..100]	Amazon Prime Video proprietary VQA model - EQM NR [11]
9	EQMfr	Integer [0..100]	Amazon Prime Video proprietary VQA model - EQM FR [11]
10	Bitmovin VQA Metrics	Integer [0..100]	Bitmovin proprietary VQA model
11	VMAFDRE	Integer [0..100]	Harmonic proprietary VQA model using DRE [12][13]
12	XVSnr	Integer [0..100]	IMAX proprietary VQA model - NR XVS [14]
13	XVSfr	Integer [0..100]	IMAX proprietary VQA model - FR XVS using EPS [14]

14	XVSeps	Integer [0..100]	IMAX proprietary VQA model - XEPS [14]
15	XVSbs	Integer [0..100]	IMAX proprietary VQA model - Banding Score
16	XVScvs	Integer [0..100]	IMAX proprietary VQA model - CVS
17	pVMAF	Integer [0..100]	Synamedia proprietary VQA model pVMAF [15]
18	MQCS	Integer [0..100]	AWS Elemental proprietary VQA model MQCS (Media Confidence Quality Score)

4. Header Usage Examples

Informative VQA-score header examples for different use cases are given below. Note that only a single CMCD-Static header is included with each CMSD response.

4.1 Single type and single GOP per segment

1. CMSD-Static: vqat="VMAF",vqas=81
2. CMSD-Static: vqat="VMAFMobile",vqas=83
3. CMSD-Static: vqat="PSNR",vqas=38
4. CMSD-Static: vqat="pVMAF",vqas=85

4.2 Single types and multiple GOP per segment

1. CMSD-Static: vqat="VMAF",vqas=(81 80)
2. CMSD-Static: vqat="VMAFMobile",vqs=(83 82 85)

4.3 Multiple types and single GOP per segment

1. CMSD-Static: vqat=("SSIM" "PSNR"),vqas=(83 38)
2. CMSD-Static: vqat=("VMAF" "PSNR"),vqas=(96 38)

4.4 Multiple types and multiple GOP per segment

1. CMSD-Static: vqat=("SSIM" "PSNR"),vqas=(83 82 38 39)
2. CMSD-Static: vqat=("VMAF" "PSNR"),vqas=(96 96 95 38 38 37)

4.5 VQA data combined with other CMSD data

1. CMSD-Static:ot=v,sf=h,st=v,d=6006,vqat="VMAF",vqas=81,br=1450,n="OriginA"

   CMSD-Dynamic: "CDNB-3ak1";etp=115000;rtt=8

References

1. Dr. Urvashi Pal, "Client-VQA for OTT ABR Streaming", *SMPTE-MetExpo, Sydney, March 2024.*
2. Dr. Urvashi Pal, "Reference-based Video Quality Analysis-as-a-Service for OTT streaming", *Submitted in the United States Patent & Trademark Office, October 2023.*
3. CTA-5006 - *"Common Media Server Data (CMSD)" , November 2022*
4. CTA-5004 - *"Common Media Client Data (CMCD)" , September 2020*
5. M. Nottingham, P-H. Kamp, "RFC 8941 Structured Field Values for HTTP", *IETF, February 2021.*
6. Zhi Li, Anne Aaron, Ioannis Katsavounidis, Anush Moorthy and Megha Manohara, "Toward A Practical Perceptual Video Quality Metric", *Netflix TechBlog, June 2016.*
7. Netflix VMAF Models*, Github Public Repository, November 2023.*
8. Zhi Li, "A VMAF Model for 4K", *VQEG Madrid 2018.*
9. Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE transactions on image processing, 2004.*
10. Z. Wang, E. Simoncelli, and A. Bovik, "Multiscale structural similarity for image quality assessment," *Thirty-Seventh Asilomar Conference on Signals, Systems & Computers, 2003.*
11. < *pph@amazon.com to confirm reference*> Zaixi Shang, Yixu Chen, Yongjun Wu, Hai Wei, Sriram Sethuraman, “Subjective and Objective Video Quality Assessment of High Dynamic Range Sports Content”, *Amazon Prime Video, Seattle, WA.*
12. Xavier Ducloux, Patrick Gendron, Thierry Fautier, “Improving Streaming Quality and Bitrate Efficiency with Dynamic Resolution Selection”, Harmonic Inc.
13. Xavier Ducloux, Jean-Louis Diascorn and Thierry Fautier, “Exploring the benefits of Dynamic Resolution Encoding and support in DVB standards”, Harmonic Inc, IBC, RAI, Amsterdam, 2022
14. < *arehman@imax.com to confirm reference* > “Imax StreamAware| On-air”, Imax Corp
15. Jan De Cock. 2023, “Video Quality Measurement & Control for Live Encoding”, *2nd Mile-High Video Conference (MHV '23). Association for Computing Machinery, New York, NY, USA, 159*