ETSI TR 103 488 V1.1.1 (2019-01)

**Speech and Multimedia Transmission Quality (STQ);
Guidelines on OTT Video Streaming;
Service Quality Evaluation Procedures**

Reference

DTR/STQ-00215m

Keywords

3G, data, GSM, network, QoE, QoS, service, video

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - NAF 742 C
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° 7803/88

Important notice

The present document can be downloaded from:
http://www.etsi.org/standards-search

The present document may be made available in electronic versions and/or in print. The content of any electronic and/or print versions of the present document shall not be modified without the prior written authorization of ETSI. In case of any existing or perceived difference in contents between such versions and/or in print, the prevailing version of an ETSI deliverable is the one made publicly available in PDF format at www.etsi.org/deliver.

Users of the present document should be aware that the document may be subject to revision or change of status. Information on the current status of this and other ETSI documents is available at
https://portal.etsi.org/TB/ETSIDeliverableStatus.aspx

If you find errors in the present document, please send your comment to one of the following services:
https://portal.etsi.org/People/CommiteeSupportStaff.aspx

Contents

Intellectual Property Rights

Foreword

This Technical Report (TR) has been produced by ETSI Technical Committee Speech and multimedia Transmission Quality (STQ).

Modal verbs terminology

In the present document "**should**", "**should not**", "**may**", "**need not**", "**will**", "**will not**", "**can**" and "**cannot**" are to be interpreted as described in clause 3.2 of the ETSI Drafting Rules (Verbal forms for the expression of provisions).

"**must**" and "**must not**" are **NOT** allowed in ETSI deliverables except when used in direct citation.

Executive summary

The fast increase of the variety, technology complexity, and dynamic changes of the OTT video streaming services' delivery, as well as their consumptions by the users on a wide range of smart devices, suggests a consistent and robust testing approach that can allow an application transparent and meaningful evaluation of these services' quality management and control. The present document offers guidance for such a testing approach, while aligning the content with all other ETSI WG STQ Mobile documents as well as work related to the topic ongoing in other organizations.

Introduction

Several published STQM TSs and TRs cover the area of the video streaming end to end QoS/QoE evaluation. ETSI TS 102 250-2 [i.1] covers video streaming session performance evaluation in the case of real time streaming (e.g. RTP, RTSP, RTCP). ETSI TR 101 578 [i.2] complements ETSI TS 102 250-2 [i.1] with QoS/QoE for the specific case of TCP based streaming (YouTube™ application), which is the de-facto delivery technique for the OTT video streaming applications. Last, but not least, ETSI TR 102 493 [i.3] provides guidance on available QoE algorithms that can be used in various video streaming testing scenarios.

However, the OTT video streaming services' arena has seen video technology evolution (e.g. H.265 codec, 1440 resolution, new protocols such as QUIC), as well as dynamic technology changes (e.g. encryption schemes, encoding schemes of various profiles, adaptations schemes at the server and/or client side within the context of various throttling techniques and policies). Last, but not least, the variety of devices, namely operation system based video clients, do see the same dynamic change. Therefore, guidance for a transparent and flexible testing approach is required by today's OTT video streaming sessions' quality evaluation. And, consequently, a set of the most meaningful QoE centric QoS parameters is also necessary, which can complement MOS estimators (QoE algorithms/models; Recommendations ITU-T J series [i.4] to [i.9], P series [i.10] and [i.11]) whenever these are available, but also go beyond a single quality score.

1 Scope

The present document's scope is to provide guidance on OTT video streaming testing approach with a set of minimum desired and most meaningful QoE centric QoS parameters along with recommendations to create a figure of merit quantifying the OTT video streaming session quality, where possible. In addition, the set of introduced QoE centric QoS parameters aim to help with the identification of the possible roots of video quality degradation. The present document also offers means to understand aspects related with network and services optimization and troubleshooting, such as the trade-off between bandwidth usage or controlled throttling and end-to-end video quality.

The scope of the present document complements ETSI TS 102 250-2 [i.1] and ETSI TR 101 578 [i.2] while not being as exhaustive, but rather focused on QoE centric characterization and an end-to-end view on the video streaming session as a whole. Furthermore, the present document takes into consideration QoE centric evaluation by means of passive, non-intrusive network monitoring of SSL/QUIC OTT Video Services bitstreams. In addition, the present document aims to complement the scope of ETSI TR 102 493 [i.3] with respect to QoE models for video streaming integrity as perceived by users.

2 References

2.1 Normative references

Normative references are not applicable in the present document.

2.2 Informative references

References are either specific (identified by date of publication and/or edition number or version number) or non-specific. For specific references, only the cited version applies. For non-specific references, the latest version of the referenced document (including any amendments) applies.

NOTE: While any hyperlinks included in this clause were valid at the time of publication, ETSI cannot guarantee their long term validity.

The following referenced documents are not necessary for the application of the present document but they assist the user with regard to a particular subject area.

[i.1] ETSI TS 102 250-2: "Speech and multimedia Transmission Quality (STQ); QoS aspects for popular services in mobile networks; Part 2: Definition of Quality of Service parameters and their computation".

[i.2] ETSI TR 101 578: "Speech and multimedia Transmission Quality (STQ); QoS aspects of TCP-based video services like YouTube™".

[i.3] ETSI TR 102 493: "Speech and multimedia Transmission Quality (STQ); Guidelines for the use of Video Quality Algorithms for Mobile Applications".

[i.4] Recommendation ITU-T P.1201: "Parametric non-intrusive assessment of audiovisual media streaming quality".

[i.5] Recommendation ITU-T P.1202: "Parametric non-intrusive bitstream assessment of video media streaming quality".

[i.6] Recommendation ITU-T P.1203: "Parametric bitstream-based quality assessment of progressive download and adaptive audiovisual streaming services over reliable transport".

[i.7] Recommendation ITU-T P.343: "Hybrid perceptual bitstream models for objective video quality measurements".

[i.8] Recommendation ITU-T P.341: "Objective perceptual multimedia video quality measurement of HDTV for digital cable television in the presence of a full reference".

[i.9] Recommendation ITU-T P.342: "Objective multimedia video quality measurement of HDTV for digital cable television in the presence of a reduced reference signal".

[i.10] Recommendation ITU-T J.247: "Objective perceptual multimedia video quality measurement in the presence of a full reference".

[i.11] Recommendation ITU-T J.246: "Perceptual visual quality measurement techniques for multimedia services over digital cable television networks in the presence of a reduced bandwidth reference".

[i.12] Recommendation ITU-T P.1401: "Methods, metrics and procedures for statistical evaluation, qualification and comparison of objective quality prediction models".

[i.13] Larry Stephens: "Schaum's Outline of Statistics" series, McGraw-Hill Trade, January 1989.

3 Definition of terms, symbols and abbreviations

3.1 Terms

Void.

3.2 Symbols

Void.

3.3 Abbreviations

For the purposes of the present document, the following abbreviations apply:

API	Application Programming Interface
CDN	Content Delivery Network
DASH	Dynamic Adaptive Streaming over HTTP
DL	DownLoad
FoM	Figure of Merit
IP	Internet Protocol
MOS	Mean Opinion Score
OS	Operating System
OTT	Over The Top
QoE	Quality of Experience
QoS	Quality of Services
QUIC	Quick UDP Internet Connections
RCA	Root Cause Analysis
RTCP	Real Time Control Protocol
RTP	Real Time Protocol
RTSP	Real Time Streaming Protocol
SSL	Secure Sockets Layer
TCP	Transport Control Protocol
TH	THreshold
UDP	User Datagram Protocol
UE	User Equipment
WG	Working Group

4 Passive, non-intrusive network monitoring of SSL/QUIC OTT video services bitstreams

OTT Video payloads transported by SSL and QUIC, together with metadata encryption, reduce the applicability of QoE algorithms/models referred by ETSI TR 102 493 [i.3] and Recommendation ITU-T J series, P series [i.4], [i.5], [i.6], [i.7], [i.8], [i.9], [i.10] and [i.11] for non-intrusive network monitoring at mid-point solutions and prevent them to provide the MOS identified by such standards.

The approach described in the present document is based on the figure of merit and on identification of the possible causes of video quality degradation, complements the one depicted by ETSI TR 102 493 [i.3] for non-intrusive network monitoring solutions and provides a reference for them when payloads are transported by SSL and QUIC.

Note that such approach can be also extended to all the passive monitoring solutions, including any possible ones on the User Equipment, when only transported payloads by SSL and QUIC are available.

5 Categories of OTT video streaming session QoS parameters

5.1 Introduction

The whole end-to-end OTT video streaming session quality is determined and impacted by four categories of QoS parameters:

- Transport/delivery QoS parameters.
- Audio/video integrity QoS parameters.
- Streaming session QoS parameters.
- Service centric QoS parameters.

Each category can cover a set of QoS parameters which can be either measured and/or calculated directly, or can be estimated based on inferences (that are out of the scope of the present document) made on client's events, service's states and transport packet analysis. The latter is needed due to various types and levels of encryptions. At the time of publication of the present document, there are scenarios within which even these inferences techniques are challenged by fully encrypted streams, such as QUIC protocol, which leaves no metadata available. Work is going on in 3GPP, IETF and DASH Industry Forum in order to offer an open API interface which would allow access to minimum required metadata.

Table 1 provides a list of minimum required and most meaningful QoS parameters which can be used for end-to-end characterization (e.g. CDNs Id, CDN protocol, subscriber IP address, session id, video provider) and quantification (e.g. video-audio QoS parameters, streaming QoS parameters) of the OTT video streaming session' s quality.Ultimately, these QoS parameters can be further used to troubleshoot and identify more likely root causes of possible quality problems.

The columns in table 1 refer to:

- **Category:** Transport; audio/video; streaming session; Service centric.
- **Name:** It provides the parameter name.
- **Type:** It refers to QoS or Auxiliary (Aux) aspects.

Table 1: QoS parameters

Category	Name	Description	Type
Transport	Number of CDN Media Servers	Number of participating CDN Media Server inside a session.	Aux
	CDN Media ServerX Name	Name of the used CDN Media Server X. X is an integer number that ranges from 1 to Number of CDN Media Servers.	Aux
	CDN Media ServerX IP address	IP Address of the used CDN Media Server X. X is an integer number that ranges from 1 to Number of CDN Media Servers.	Aux
	CDN Media ServerX downloaded bytes	Bytes downloaded from CDN Media server X X is an integer number that ranges from 1 to Number of CDN Media Servers.	Aux
	CDN Media ServerX Average Time to first packet (CDN delay)	Average time between the last packet of the player request till the first packet of the related CDN Media server response for all the player requests with a response. X is an integer number that ranges from 1 to Number of CDN Media Servers.	QoS
	CDN Media ServerX Failure rate (%)	Rate of Player Requests with no answer towards CDN Media ServerX. X is an integer number that ranges from 1 to Number of CDN Media Servers.	QoS
	CDN Downlink Application throughput (Kbps)	It is the overall Downlink Application Throughput related to media content downloaded from CDN Media Servers (Kbps).	QoS
Video/Audio	Avg. video buffer (s)	Average Player video buffer size during play time in second.	Aux
	Avg. audio buffer (s)	Average Player Audio buffer size during play time in second.	Aux
	Avg. video bit rate (Kbps)	Average video bit rate during play time (Kbps).	QoS
	Avg. video bit rate Range	Average video bit rate range during play time (low, fair, excellent). See clause 6 for Range definition.	QoS
	Avg. audio bit rate Range	Avg. audio bit rate during play time (low, fair, excellent). See clause 6 for range definition.	QoS
	Low Video bit rate %	Play time % at Low Video Bit Rate range. See clause 6 for range definition.	QoS
	Low Audio bit rate %	Play time % at Low Audio Bit Rate range. See clause 6 for range definition.	QoS
	Good Video bit rate %	Play time % at Good Video Bit Rate range. See clause 6 for range definition.	QoS
	Good Audio bit rate %	Play time % at Good Audio Bit Rate range. See clause 6 for range definition.	QoS
	Excellent Video bit rate %	Play time % at Excellent Video bit rate range. See clause 6 for range definition.	QoS
	Excellent Audio bit rate %	Play time % at Excellent Audio bit rate range. See clause 6 for range definition.	QoS
	Video bit rate range switches	Number of Video bit rate range switches. See clause 6 for range definition.	QoS
	Audio bit rate range switches	Number of Audio bit rate range switches. See clause 6 for range definition.	QoS
	Positive Video bit rate range switches	Number of switches to higher video bit rate range. See Clause 6 for range definition.	QoS
	Positive Audio bit rate range switches	Number of switches to higher audio bit rate range. See clause 6 for range definition.	QoS
	Negative Video bit rate range switches	Number of switches to lower video bit rate range. See clause 6 for range definition.	QoS
	Negative Audio bit rate range switches	Number of switches to lower audio bit rate range. See clause 6 for range definition.	QoS
	Video DL Mbytes	Video downloaded Mbytes.	Aux
	Audio DL Mbytes	Audio downloaded Mbytes.	Aux
	Video Quality	It refers to ETSI TR 101 578 [i.2] Video Quality parameter.	QoS

Category	Name	Description	Type
streaming session	Video Access Time (s)	It refers to ETSI TR 101 578 [i.2] Video Access Time (s) parameter, see also clause 5.1.	QoS
	Pre-Playout Buffering time (s)	It refers to ETSI TR 101 578 [i.2] Pre-Playout Buffering time (s) parameter, see also clause 5.1.	
	Video Playout Duration (s)	It refers to ETSI TR 101 578 [i.2] Video Playout Duration (s) parameter.	QoS
	Number of freezes	Video Freeze Occurrences.	QoS
	Accumulated Video Freezing duration (s)	It refers to ETSI TR 101 578 [i.2] Accumulated Video Freezing Duration (s) parameter.	QoS
	Video Freezing Time proportion	It refers to ETSI TR 101 578 [i.2] Video Freezing Time Proportion parameter.	QoS
	Video Playout cut-off	It refers to ETSI TR 101 578 [i.2] Video Playout cut-off parameter.	QoS
	Terminated during freeze	In addition to ETSI TR 101 578 [i.2] Video Playout cut-off, it indicates if the video was terminated during a freeze (0 if no, 1 if yes).	QoS
	Video Access Failure	It refers to ETSI TR 101 578 [i.2] Video Access Failure parameter.	QoS
	Terminated before video reproduction start	Differently from ETSI TR 101 578 [i.2] Video Access Failure or Pre-Playout buffering failure, it does not distinguish if the termination was due to the subscriber that exits before the video start or it was due to a player exit due to bad network condition.	Aux
	Video transfer time (s)	Time from the first to the last data packet from CDNs in second.	Aux
	Starving time (s)	Time in second with no responses to player requests.	Aux
Service centric	Subscriber IP address	UE IP address.	Aux
	Video provider	YouTube™, Netflix™, etc.	QoS
	OS Client	OS client: Android™, iOS™, etc.	Aux
	Player Type	browser, App.	Aux
	CDN transport protocol	HTTP, QUIC, SSL, etc.	Aux

Parameters that are common with ETSI TR 101 578 [i.2] should use the timeout specified therein.

5.2 Triggers of main QoS parameters

5.2.1 Session

A video streaming session is the interval time during which a subscriber executed one or more consecutive videos from the same OTT provider.

A session should start when the first contact from the player is detected towards a CDN Media server to play one or more consecutive videos. It should end after an inactivity time between the Player and the CDN Media Servers.

5.2.2 Video Access Time and Pre-playout buffering time

Monitoring solutions based on player API observation should characterize the video access phase from the video clip request till the playout start as described in figure 1.

Non-intrusive passive monitoring solutions should evaluate the Pre-playout buffering time portion of the Video Access Time, being the Video clip request times not observable because of SSL and QUIC encryption.

5.3 Parameters relevance for an effective Root Cause Analysis

5.3.1 Transport parameters relevance

CDN QoS parameters allow to identify the involved media servers and enable to detect the ones related to low QoE and to address a further analysis about them.

"CDN downlink application throughput" parameter characterizes the connectivity performances experienced during the video reproduction and allows to identify low QoE related to it.

5.3.2 QoE-Video/Audio parameters relevance

Avg. video/audio buffer allows an historical trend analysis on the buffer size strategy adopted by the different players. In line of principle, lower this number, higher the probability that freezes are experienced.

Video/Audio DL Mbyte enable to get a view of the network usage in term of volume. When it is referred to different network elements, it could identify possible bottleneck related to them.

The value of Low, Good, Excellent Audio/video bit rate % parameters is explained in clause 6.

Audio/Video switches indicators could contribute to detect sessions with a high number of switches that lower the streaming service's quality from the user's perspective point of view.

5.3.3 Streaming session parameters relevance

Long "Video Access Time" or "Pre-Playout Buffering Time", "Freezes", "Accumulated Freeze duration", "Video Freeze Time Proportion", "Terminating During Freeze" identify events that lower user experience. Therefore, they should be the main driver to start Root Cause Analysis (RCA). The interval time when such event happened could also be relevant for RCA. They could identify Critical Interval Times when the user experience was not good. Common Critical Interval Times between different video sessions could detect problems affecting more users.

High concentrations of "Terminated Before Video Reproduction Start"/"Video Access Failure" and "Starving Time" in an interval time when Video Access Phase/Pre-Playout buffering time are high or Freezes are present could be related to a more general issue than lowered QoS due to a main problem on the transport network or on the CDN side. By analyzing the distribution of all these parameters by "Video provider", it should be possible to address RCA towards the operator network or towards the CDN Media server side, as explained in clause 5.3.4.

"Video Transfer Time" analysis, when aggregated by CDN, should be useful to detect CDN Media server utilization and identify overload of them that result on low QoE parameters (Long Video Access Time/Pre-Playout Buffering Time or Freezes).

5.3.4 Service centric parameters relevance

"Video provider" allows to get different FoM for different providers and to benchmark them too. If, in the same interval times, low QoS is detected in all the video provider, this may suggest an issue in the operator network that delivers the service. If it is more related to a specific OTT video provider, that may suggest an issue on the CDN Media server side.

Player and OS type enable to address RCA of low QoS streaming session parameters to specific OS (Android™, iOS™, etc.) or Player types (Browser, App).

"Transport Protocol" allows to aggregate the QoS Audio/Video and streaming session parameters to compare performances of the video delivered by QUIC, SSL and HTTP, and to address specific Root Cause Analysis on the network elements related to them.

6 Thresholds for OTT video streaming quality

In order to use the Qos parameters (table 1) to identify quality concerns and/or optimization needs, and then act upon the findings, the Qos parameters values need to be compared against pre-defined quality thresholds.

Table 1 shows three thresholds for the video-audio bit rates: low, good and excellent. Their scope is to quantify video quality, respectively video streaming session's integrity, based on the video bit rate, one of the significant factors impacting the video integrity QoS and therefore the first one to look at in order to understand video quality behaviour. Only three thresholds are used, since for mobile applications these three thresholds mainly trigger either quality concerns (e.g. below good) and/or real need for optimization and/or control of throttling schemes (e.g. below low). The "excellent" threshold is generally used more or less as an indicator for no quality concerns or even as an indicator for the possibility to increase capacity; very likely more video streaming sessions would be possible at good quality if "excellent" threshold is consistently met. The selection of the values of these thresholds should be decided based on generally expected bit rates depending on the OTT video application, used codec/container/profile, etc. It is important to bear in mind that these thresholds are expected to change once higher resolutions or new codes are available, supported by mobile devices and fairly spread out.

Quality thresholds for other QoS metrics (e.g. buffering, play delay, freezing time) can be pre-defined based on video application type as well as operator's QoS polices. For example, a YouTube™ clip of 2 min length and showing an initial buffering time of more than 30 sec is more annoying for a user's perception than a Netflix™ video of 30 min length with the same 30 sec initial buffering time.

In addition, operators might choose to decide when to trigger quality concerns and/or optimization needs as well as what should be the pre-defined quality thresholds for various QoS parameters based on their own business interests and network policies and strategies.

As an example, the present document shows some default Video and Audio Bit Rate thresholds settings based on current status of the technology, streaming techniques/protocols, codecs/players, resolutions. Note that if the "video manifest" is aware that the player downloads before to start the reproduction, such thresholds can be setup according to the information provided to it for each video, thus improving the matching between the video bit rate and the related video resolution. When the "video manifest" is not accessible (like when only QUIC and SSL packets are analysed) Video Bit Rate/Resolution mapping can be based on a classification thresholds approach which make assumptions on the video characteristics like the ones used on the following examples. Moreover, each OTT video service could have its own mapping, once characterized by the used codecs.

Table 2: Video Bit Rate thresholds _ default settings assuming H.264, 24-30 fps, Dynamic (non static) videos

	Lower limit (≥)	Upper limit (<)
Low Video Bit Rate Range (≤ 480 p)		1 400 Kbps
Good Video Bit Rate Range (720 p)	1 400 Kbps	3 200 Kbps
Excellent Video Bit Rate Range (≥ 1080 p)	3 200 Kbps	

Table 3: Audio Bit Rate thresholds _ default settings

	Lower limit (≥)	Upper limit (<)
Low Audio Bit Rate Range		55 Kbps
Good Audio Bit Rate Range	55 Kbps	100 Kbps
Excellent Audio Bit Rate Range	100 Kbps	

7 Figure of Merit of the OTT video streaming service

7.1 Introduction

In order to quantify the OTT video streaming service quality as perceived by users, a Figure of Merit can be defined.

Table 1 contains a required set of QoS metrics which can be used to characterize, quantify and consequently trigger quality concerns and/or optimization needs.

However, a Figure of Merit (FoM) describing the overall video streaming service quality is not required to use all the QoS parameters presented in table 1, but rather the main contributors to the overall streaming service's quality from the user's perspective point of view. As an example, Figure 1 presents the YouTube™ video streaming case as described in ETSI TR 101 578 [i.2]. The QoS parameters determining and impacting the perceived waiting time, the perceived video media quality (integrity) and the perceived retainability are those recommended for the calculation of the figure of merit of the overall video streaming service quality. It should be noted that unlike all the ITU-T video QoE metrics (e.g. Recommendations ITU-T P.1201 [i.4], P.1202 [i.5], P.1203 [i.6] series, P.343 [i.7] series, P.341 [i.8], J.342 [i.9], J.247 [i.10] and J.246 [i.11], which refer only to the video sessions integrity, the FoM refers to the performance of the whole video streaming session, from access, to integrity of the media and to retainability of the session, and consequently of the whole service.

However, in some encrypted scenarios (e.g. QUIC), and depending as well on test use case (e.g. passive bit stream monitoring or on-device monitoring through APIs), some of the video phases and trigger points are not fully observable and/or accessible (e.g. black dotted lines in figure1). In addition, other video streaming applications might not display all the described sub-intervals characteristic showed in figure 1. In these cases, the FoM can only partially describe the overall video streaming service quality, in terms of waiting time, integrity and retainability. If a FoM is calculated in these cases, then it is mandatory to be reported along with its validity and limitations.

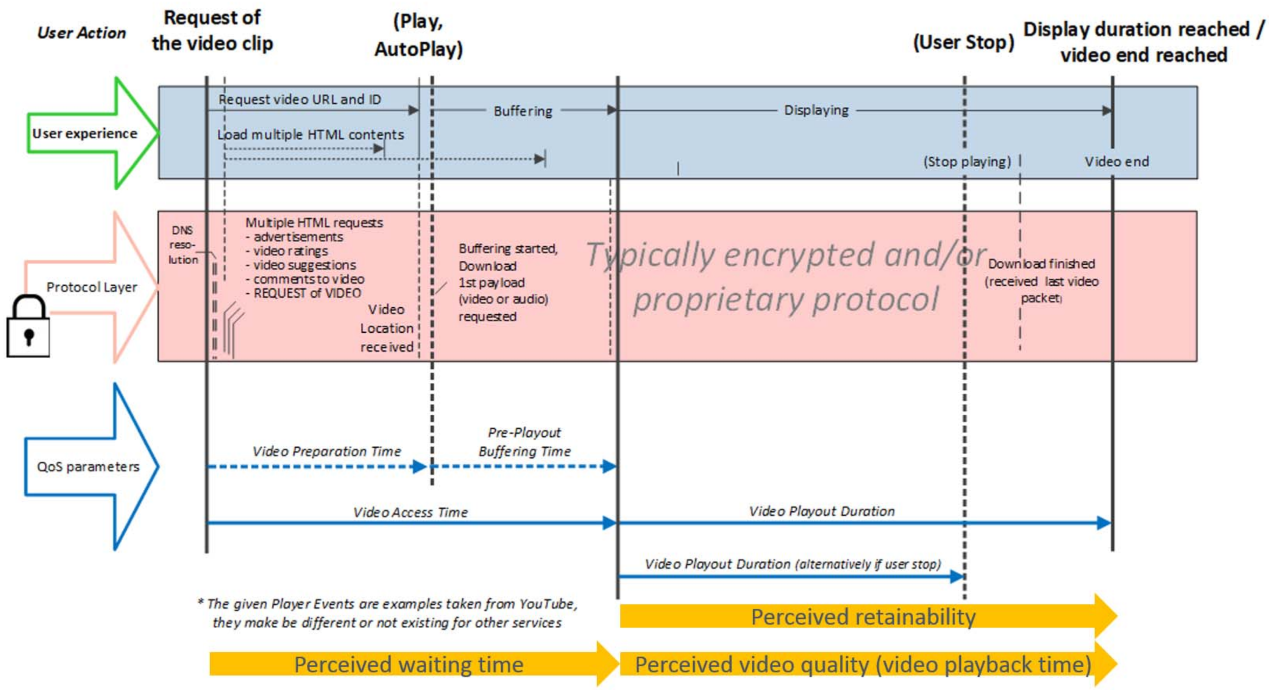

Figure 1: Typical phases of IP-based video services

7.2 Calculation of Figure of Merit

The Figure of Merit is determined as a statistical score (see Recommendation ITU-T P.1401 [i.12]) defined by all QoE centric QoS_i (i = 1,N) metrics affecting the overall perceived quality of the video streaming service. Thus, the FoM is defined by a weighted sum of the StatDiff_i of each QoS parameter vs. the pre-defined quality threshold QTh_i (i = 1,N):

$$\mathrm{FoM} = \Sigma(w\_i \times \mathrm{StatDiff\_i})$$

where w_i is the weight allocated to each QoS parameters contributing to the perceived service quality. As described in Recommendation ITU-T P.1401 [i.12], it is assumed that the values of each QoS metric can be represented by a Gaussian distribution; the assumption of a Gaussian distribution stands for a sample population larger than N > 30 samples [i.13]. In addition, for the comparison against the quality threshold, it is assumed that the threshold is described by the same distribution (e.g. same statistical metrics such as average and standard deviation, number of samples) as the analysed QoS metric.

The FoM expressed by the statistical score represents the statistical difference (StatDiff) versus the quality threshold. The StatDiff of each QOS_i vs. its QTh_i is defined (Recommendation ITU-T P.1401 [i.12]) by equation (1) if the QOS can be represented by a continuous random variable (e.g. play delay) and by equation (2) if it can be represented by a discrete random variable (e.g. failure/success ratios, percentages).

$$\text{StatDiff}\_i = \max\{0, (QTh\_i - QoS\_i / \text{sqrt}(2 \times std\text{^}2/N) - F(0{,}05, N, N)\} = \max\{0, Zn - F(0{,}05, N, N) \qquad (1)$$

$$\text{StatDiff}\_i = \max\{0, (QTh\_i - QoS\_i) / \text{sqrt}(2 \times p1 \times (1-p1)/N) - F(0{,}05, N, N)\} = \max\{0, Zn - F(0{,}05, N, N)\} \qquad (2)$$

Equations (1) and (2) are based on statistical significance testing (hypothesis testing) [i.13], where F(0,05, N, N) is the Fisher statistics with N degree of freedom, with 95 % confidence. N represents the number of samples of each contributing QoS_i (i = 1,N). As mentioned above, the Gaussian distribution is assumed for the continuous variables, which is statistically valid for N > 30 [i.13]. In the case of the discrete variables, those are represented by Binominal distribution, which can be approximated by a Gaussian distribution, if N >30 [i.13].

The lower the FoM, the closer to the quality thresholds of all considered QoS parameters, consequently the better the overall video streaming service quality.

7.3 Quality thresholds used for FoM calculation

The FoM describes the performance versus a pre-defined (required or desired) quality threshold. Therefore, these thresholds need to be a-priori defined, as discussed in clause 6. Some examples are presented in annex A. It should be also noted that the quality thresholds can change depending on the introduction of new technologies, new video type of streaming applications and service, etc.

In the case in which a quality range (minimum to maximum value) is preferred versus a single value, the FoM can be determined relative to both ends of the range. The lower the FoM, the closer to the minimum or maximum value of the quality range. However, in order to take straightforward decisions on the video streaming service performance based on FoM determination, it is recommended to use as quality threshold the average of the two values, minimum and maximum.

7.4 Weightings

The definition of the weightings is out of scope of the present document. These weightings could for example be defined by performance demands, business strategies and policies.

In addition, weightings could be defined based on the type of the streamed video, as briefly introduced in clause 6.

For example, for an YouTube™ video is more likely that waiting time is more important than the integrity of the media. This is because it is expected that viewing a shorter clip is subjectively affected more by how much would one need to wait to download the video rather than how good the quality of the media is, as long as it is not too degraded and it allowed the capturing of the message contained in the video. On the other hand, video services like Netflix™ are expected to be subjectively affected more by the integrity of the media rather than the waiting time to download an one hour or more of video.

7.5 Video streaming sessions' quality classification

FoM expresses the performance of the overall video streaming session, and therefore it is recommended to be calculated for all the sessions during a pre-defined time window. Therefore, it should take into consideration all session regardless if they show consistent bad and/or unexpected behaviour, such as very long freezing time or high rate of failure.

However, for a complete and thorough performance analysis of the video streaming service, it is recommended to report FoM values along with video streaming quality classes, such as, but not limited to: percentage of normal sessions, percentage of completely failed sessions, percentage of sessions with unexpected long initial buffering, percentage of sessions with very long freezing time.

Annex A presents couple of examples of how FoM of video streaming service quality can be calculated and interpreted.

8 Reporting

If the FoM is published externally, all applied parameters, particularly all thresholds and weights, should be reported together with the FoM.

Annex A: Example of FoM calculation

The following scenario is presented as an example of FoM calculation.

It is desired to evaluate the YouTube™ video streaming service quality as perceived by users based on 100 video sessions. Two possible test scenarios are considered.

The first test use case refers to a passive bit streaming monitoring solution. In this case the available QoS parameters, are (table A.1):

- Pre-Playout Buffering Time.
- Video Freezing time proportion.
- Avg Video Bit rate.
- Session Terminated during freeze.

Table A.1 shows also pre-defined quality thresholds and selected weights. It should be noted that this is just an example and that thresholds and weights should be decided based on what is presented in clauses 6 and 7 of the present document. In the weights case, it can be seen that the weights selected for perceived waiting time and retainability are higher than the ones selected for integrity. The reason is the fact that the example discusses the YouTube™ application and therefore the perception is impacted as described in clause 6.

The second test use case refers to on-device monitoring solution (APIs accessible) for which the available QoS parameters are (table A.2):

- Video Access Time.
- Video Freezing Time proportion.
- Video Quality (estimated MOS).
- Video Playout cutoff.

Table A.2 shows also pre-defined quality thresholds and selected weights. It should be noted that this is just an example and that thresholds and weights should be decided based on what is presented in clauses 6 and 7 of the present document.

As already mentioned in clause 7, the FoM can use more QoS parameters than the ones provided in these examples. However, it should be kept in mind that the goal of FoM is to provide an indicator regarding the performance of the video session as perceived by the users.

Table A.1: QoS parameters example for passive bit streaming monitoring solution

QoS parameter (impacting overall session QoE)	Required performance threshold	Weight	Perceived Quality categories
Pre-Playout buffering time (s)	≤ 5 s	0,3	Waiting time
Video Freezing time proportion (%)	< 5 %	0,2	Integrity/Playout
Avg Video Bit Rate (Kbps)	> 800 Kbps	0,2	Integrity/Playout
Sessions terminated during freeze rate (%)	≤ 0,5 %	0,3	Retainability

Table A.2: QoS parameters example for on-device monitoring solution

QoS parameter (impacting overall session QoE)	Required performance threshold	Weight	Perceived Quality categories
Video Access Time (s)	≤ 5 s	0,3	Waiting time
Video Freezing time proportion (%)	< 5 %	0,2	Integrity/Playout
Video quality (estimated MOS)	> 3,8 MOS	0,2	Integrity/Playout
Video Playout Cut Off (%)	≤ 0,5 %	0,3	Retainability

Across the 100 video sessions, the average and standard deviation values are calculated for each of the metrics; the results are presented for the two measurement cases; table 3 for passive bit streaming and table 4 for on-device scenario. In addition, by using equations (1) and (2) the StatDiff is calculated; the F function for N=100 and at 95 % level of confidence is F(0,05, 100, 100) = 0,716. By using the weights and the calculated StatDiff, the FoM is determined (see tables A.3 and A.4). In order to interpret the determined FoM, its minimum value has to be first calculated. The minimum value is achieved when each of the analysed QoS/QoE parameters meet their required performance threshold (tables A.1 and A.2), which is FoMmin = 0,719. Any deviation from this value shows network degradation. The same FoMmin is valid for the both measurement use cases since it is determined for the same quality thresholds. The lower the FoM, the closer to the requested and/or desired performance expressed by FoMmin.

The results presented in tables 3 and 4 show that using the two measurement scenarios, passive bit streaming and on-device use case, the performance of the 100 video sessions have almost the same FoM. This result is expected since both measurements have been used to evaluate the same 100 YouTube™ sessions. In addition, the contribution to the FoM (tables 3 and 4) is calculated for each QoS parameter and it represents the QoS average value multiplied by the correspondent weight (see formula in clause 7.2). It can be seen that the highest contribution in degrading the video session quality comes from the pre-playout buffering time (table A.3) for the passive bit streaming monitoring use case, and video access time (table A.4) for the on-device streaming monitoring use case, Therefore, it can be concluded that the worst performing QoS parameter impacting the overall video session quality is the pre-playout (or video access). This information should be used for root cause analysis.

Table A.3: FoM example for passive bit streaming monitoring solution

	Avg QoS parameter value	std	StatDiff	QoS_TH	Zn	weight	Contribution (w*StattDif)
Pre-Playout buffering time (s)	7	0,9	16,432	5,000	-15,71	0,3	4,929
Video Freezing time proportion (%)	0,061	0,239	1,044	0,050	-0,325	0,200	0,208
Avg.Video bit rate (kps)	0,900	0,200	4,254	0,800	-3,536	0,200	0,85
Sessions terminated during freeze rate (%)	0,010	0,099	1,074	0,005	-0,355	0,3	0,322
FoM			6,311				

Table A.4: FoM example for on-device streaming monitoring solution

	Avg QoS parameter value	std	StatDiff	QoS_TH	Zn	weight	Contribution (w*StatDiff)
Video access time (s)	8	13	17,036	5,000	-16,3178	0,3	5,11
Video Freezing time proportion (%)	0,061	0,239	1,044	0,050	-0,325	0,200	0,208
Video Quality (estimated MOS)	3,200	0,700	5,342	3,800	6,060915	0,200	1,068
Video playout cut off (%)	0,010	0,099	1,074	0,005	-0,35533	0,3	0,322
FoM			6,710				

Annex B: Bibliography

- Dash Industry Forum:

NOTE: Available at https://dashif.org/.

History

Document history		
V1.1.1	January 2019	Publication