VQEG

Key performance indicators and quality of experience

June 2025

Pablo Pérez, François Blouin, Kjell Brunnström
VQEG 5G-KPI Working Group

VQEG — A Brief History

- Formed in 1997 to advance the field of video quality assessment
- Closely related to ITU-T and ITU-R study groups
  - ITU-T SG9 (Broadband cable & TV)
  - ITU-T SG12 (Performance, QoS and QoE)
  - ITU-R WP6C (Programme production and quality assessment)
- Historically, a primary focus on:
  - Creation of test plans to develop and validate objective quality metrics
  - Particular focus on defining the scope and subjective test methodology
  - Statistical techniques for assessing model performance
  - → recommending approaches/models to be standardized by ITU-R/ITU-T

Meta, USA, 2025

INSA Rennes, 2022

Turin, Italy, 1997

Wuhan Uni., Tencent, China, 2019

Google, USA, 2019

How is it organized?

- VQEG Co-chairs
  - Kjell Brunnström (RISE Research Institutes of Sweden AB)
  - Ioannis Katsavounidis (Meta)
- Working groups
  - Individual co-chairs per group
- Bi-annual meetings
  - Historically in-person, worldwide
  - Winter online
  - Spring in-person
- Next meeting: November/December 2025


RI.SE
Meta
NETFLIX
TestDevLab
Disney
Adobe
Kingston University London
PlayStation
AT&T
YONSEI UNIVERSITY
ERICSSON
TECHNISCHE UNIVERSITÄT ILMENAU
SVT
DOLBY DIGITAL
NTT
NIPPON TELEGRAPH AND TELEPHONE CORPORATION
Google
NOKIA
UNIVERSIDAD POLITÉCNICA DE MADRID
FEDERAL MINISTRY OF COMMUNICATIONS, INNOVATION & DIGITAL ECONOMY
Innowireless
UNIVERSITÉ DE NANTES
SAMSUNG
AGH UNIVERSITY OF KRAKOW
AGH
ANATEL
UNIVERSITY OF THE WEST of SCOTLAND
UWS
London
aveq
Politecnico di Torino
1859
DIT
DEGGENDORF INSTITUTE OF TECHNOLOGY
UNIVERSITEIT GENT
UNIVERSITY OF SURREY

What's nice about VQEG?

- Free to join — no membership fees
- No strict or complicated rules
  - Consensus is often reached without lengthy voting procedures
- Simple organization and hierarchy
  - Chairs & co-chairs for different projects
  - Anyone can propose or contribute a new project
- Highly interactive meetings
  - Anyone can present their ideas
  - Focus on discussion time
- Not a commercial venue
  - No sales talks, no commercial advertising
- Mixture between academia and industry

5G Key Performance Indicators (5GKPI)

- Goals:
  - Defining use cases for video in 5G
  - Studying QoE aspects for video in mobility and industrial scenarios
  - Identifying the relevant network KPIs and application-level video KPIs (e.g. picture quality, A/V sync, …)
  - Building open datasets for algorithm testing and training
- Recent highlights:
  - ITU-T Technical Report GSTR.5GQoE (2022): Specific QoE requirements and required performance and features from the network

Current Focus: towards 6G

Current status: missing visibility

- Operators are missing visibility on content quality
- CAP lacks visibility on network status

Use cases: today and future

- Today: streaming + video calls
- Future: XR, tele-operation…

What can VQEG do?

Identify relevant QoE-QoS activities (ITU-T, VQEG, NGMM, TIP, IETF, 3GPP…)

Organize collaboration between operators – content providers – vendors

Agree on a common language for QoS and QoE in networks

Propose an approach to model QoS/QoE relation

Define a framework for QoE management

VQEG

White Paper: Quality of Experience aware management for collaboration between network and application providers

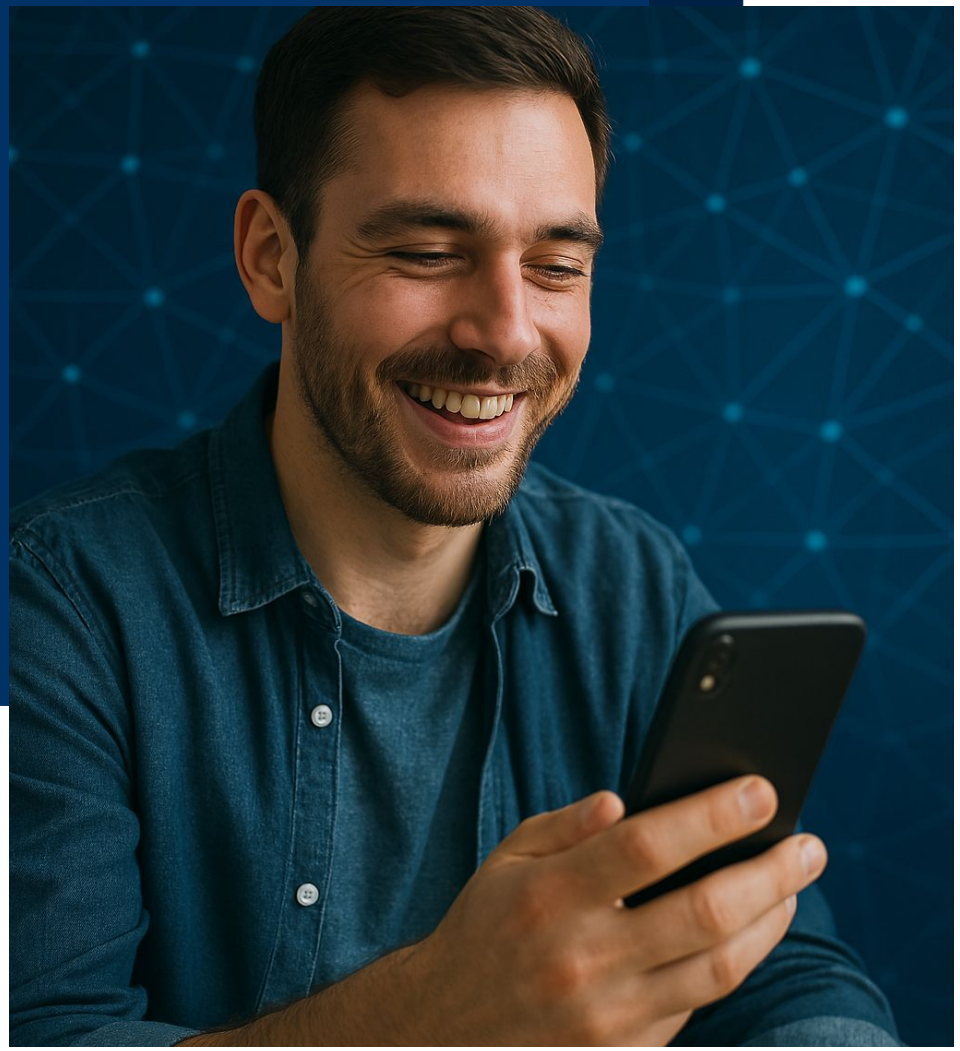
5G-KPI Working Group

Contributors

- Pablo Pérez, Pablo Rojo, Marta Orduna (Nokia)
- François Blouin, Ioannis Katsavounidis (Meta)
- Balu Adsumilli, Neil Birkbeck (Google/YouTube)
- Theo Karagioules, Emir Halepovic (AT&T)
- Jesús Folgueira, Luis Miguel Contreras (Telefonica)
- David Lindero (Ericsson)
- Werner Robitza (AVEQ)
- Kjell Brunnström (RISE)
- Michele Zorzi, Federica Battisti, Sara Baldoni (University of Padua)
- Alex Raake, Rakesh Rao Ramachandra Rao (TU Ilmenau)
- Tobias Hoßfeld (Julius-Maximilians-Universität Würzburg)
- Narciso García (Universidad Politécnica de Madrid)
- Kamil Koniuch (AGH University of Science and Technology)
- Markus Fiedler (Blekinge Institute of Technology)

Table of Contents

VQEG



01

Introduction and Scope

Towards better tools for QoE management in communication networks

VQEG

Vision

North Star

Develop QoE/QoS models and framework for QoE management in 5G-advanced/6G.

Motivation

- More efficient end-to-end video delivery
- Identify and solve the problems seen by users
- Ecosystem win-win-win approach CAP/CSP/User
- Overcome lack of trust and enhance cooperation among stakeholders

VQEG

Scope

- QoE metrics: this paper will rationalize and propose
  - Clear definition and understanding of metrics,
  - Industry alignment (CAPs, CSPs) on proposed metrics to use, share and expose,
  - Recommendation on standardization framework
  - How to measure and action.
- Initial focus on CAP / CSP interaction
  - Acknowledging that other stakeholders exist
- Generic methodology , but applied to specific use cases
  - Start: multimedia streaming and communication over a mobile network
  - At least 2 services (CAPs) and 2 network types (CSPs)
  - Framework for information exchange between CAPs and CSPs

VQEG

Towards a win–win approach (High walls between CAPs and CSPs)

CSP CHALLENGES

- Lack of access to QoE and traffic type → Inference – prone to inaccuracies and hardly actionable
- Diverse QoE targets and QoS requirements per traffic class and per wireless product (mobile vs. fixed)
- Reconciling best effort delivery with different fairness, QoE, requirement and cost profiles of apps
- Dealing with new traffic types and services

CAP CHALLENGES

- Identifying network health, congestion, available bandwidth, loss, latency
- Network policies to determine optimal streaming profile and adaptation behavior
- Opacity to QoS impairments
- E2E view to perform QoE optimization

Lack of KPI uniformity or standards among CAPs, CSPs and between CAPs and CSPs

VQEG

Issues & Gaps

CSP Perspective

- **Diverse Requirements**:
  - Complexity in policy design, mapping QoS fairness to QoE fairness.
- **Lack of QoE Information:**
  - Limits utility of QoS frameworks (e.g., 3GPP QCI), relies on inaccurate inference.
- **Lack of Real-Time QoE Feedback:**
  - Missed opportunities for proactive optimization beyond basic KPIs.

Opportunities via Exchange: Targeted QoS/QoE policies, accurate traffic handling, informed network monitoring/tuning , real-time QoE-aware actions.

CAP Perspective

- **Determining Network State:**
  - Capacity/Congestion leads to stalls, poor ABR choices. Content adaptation for interactive gaming encoding
- **Predicting Network Dynamics:**
  - Variability can create stalls, retransmissions and data waste.
- **Understanding Subscriber Context:**
  - Data plans/limits leads to potential overage or unnecessary throttling.

Opportunities via exchange: Improved bandwidth estimation, better ABR tuning, optimized prefetching/buffering, efficient data usage.

VQEG

Intro to QoE

Quality of Experience (QoE)

Degree of delight or annoyance of the user of an application product or service such as audio/video fidelity in a video session, relative to their expectations

Product Metrics

e.g, user count, user retention, user watch time

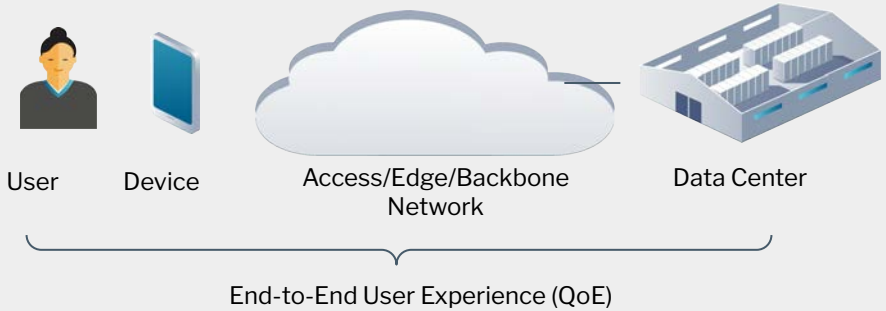
QoE User and Application Metrics

subjective impressions measured by metrics such as Mean Opinion Score (MoS),or satisfaction ratings
e.g. video/voice quality, rebuffering, stalls

Quality of Service (QoS)

Totality of characteristics of a network infrastructure and QoS mechanisms that bear on its ability to satisfy stated and implied needs of the user of the service

QoS Infrastructure Metrics



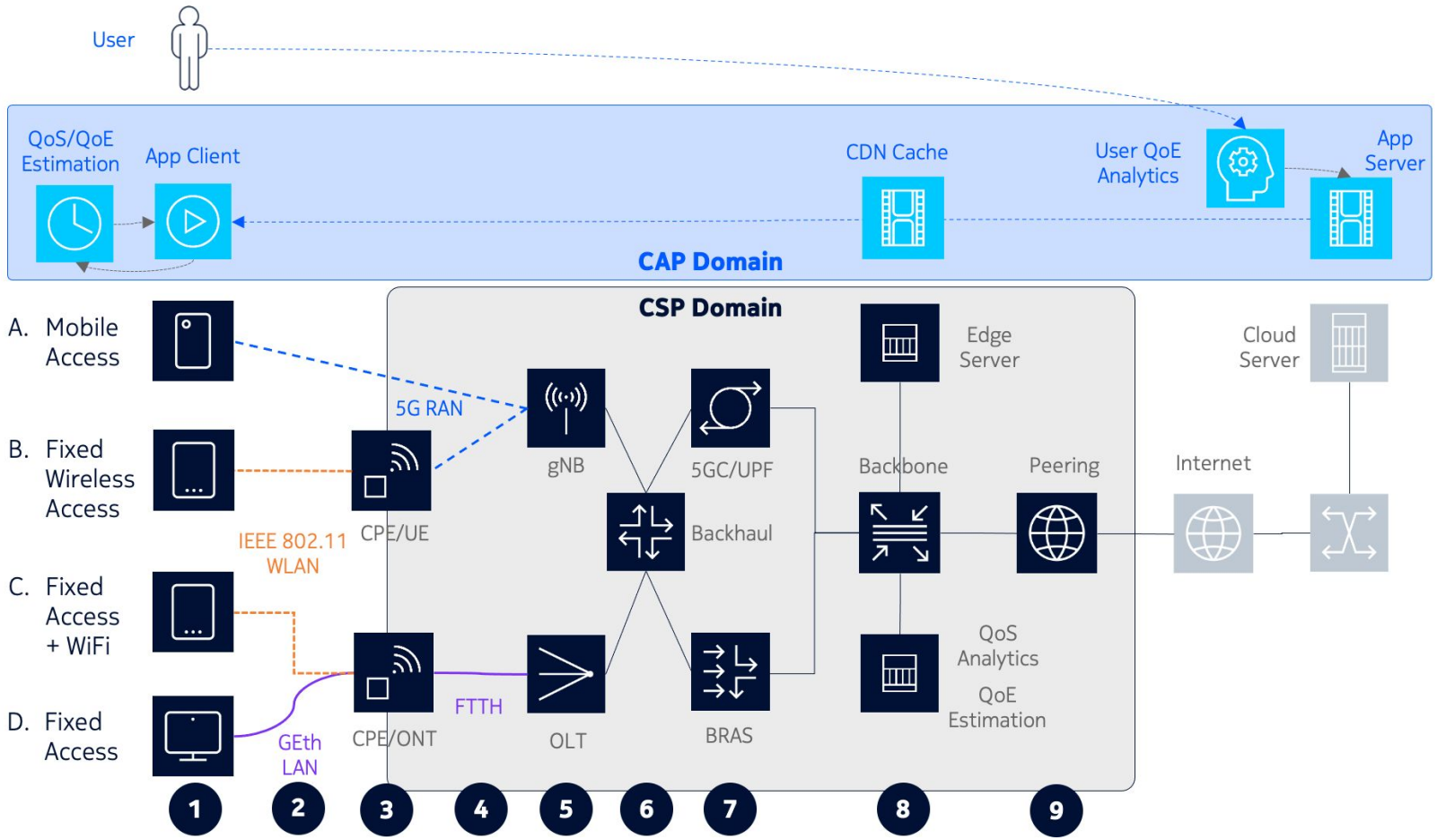
e.g., throughput, latency, loss, jitter, congestion

"QUALINET White Paper on Definitions of Immersive Media Experience (IMEx)", QUALINET meeting (online), 2020.

Use cases

Category	Type	Example
Traditional 3GPP Applications	Messaging	SMS and MMS
	Voice	CS Voice, VoLTE, VoNR and VoWIFI
Connected Application	Web Browsing	News, eLearning, shopping
	Social Media	X, Facebook, Instagram, Tiktok
	Internet of Things	Wearables, Connected home
	Online Gaming	Massively multiplayer online games
Real-time Communications (RTC)	Video Conferencing	WhatsApp,, Zoom, Teams, etc.
	Voice OTT	
	Messaging	
Streaming Media	Long Form Media	Netflix, Spotify, YouTube, etc.
	Short Form Media	TikTok, Instagram Reels, YouTube Short
	Live Media	Live Events (Sport, Music, etc.)
Immersive Applications	Cloud Gaming	Nvidia, Xbox, Luna, PlayStation, etc.
	xR (extended reality)	Industrial Applications
File Transfer	Backup, Upgrades and Transfers	Dropbox, Windows Update, SpeedTest, etc.

Network architecture landscape



Discussion

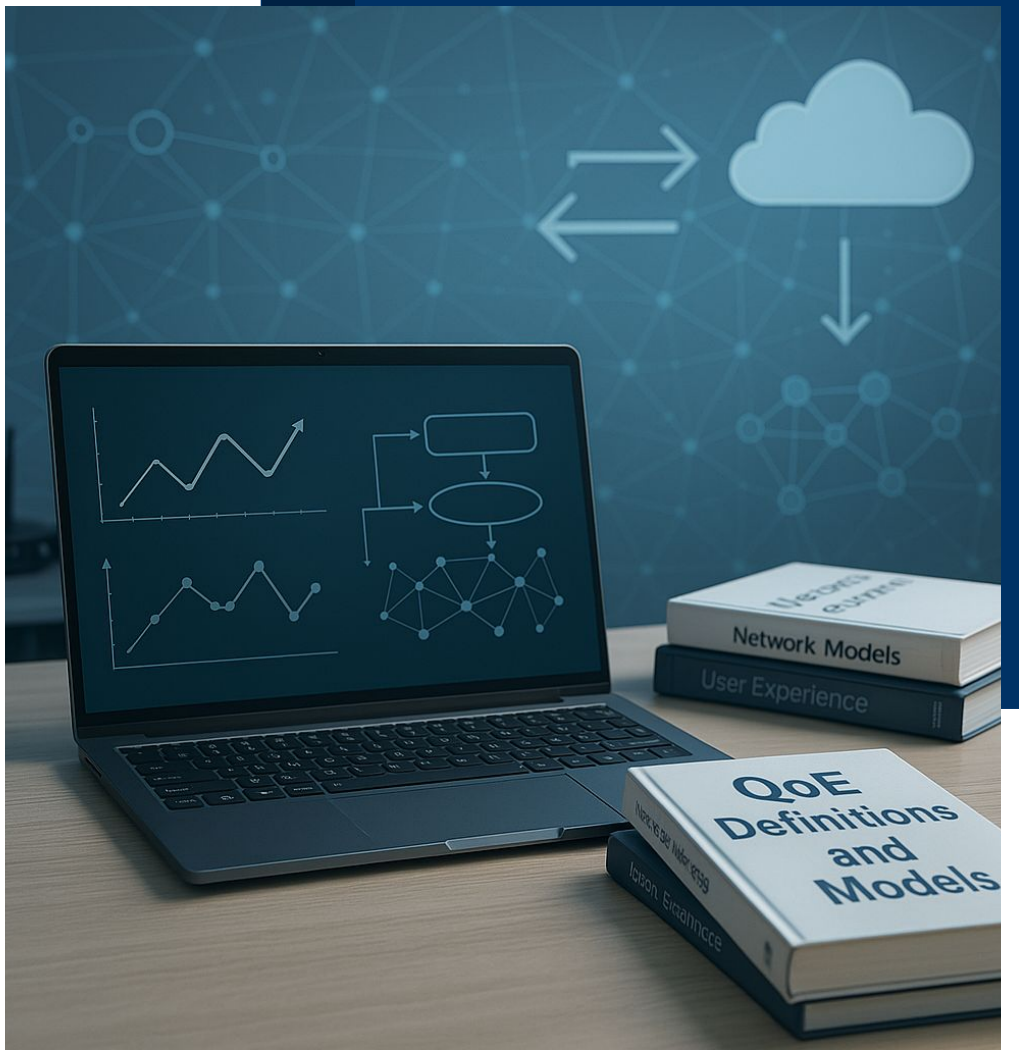
- In the interaction between CAPs and CSPs to improve QoE
  - Which are the most important challenges?
  - Is there any relevant issue that is not being addressed in the scope of the white paper?

- Which are the relevant services / use cases?

- Where are the network bottlenecks?

QoE Definitions and Models

Theory and best practices

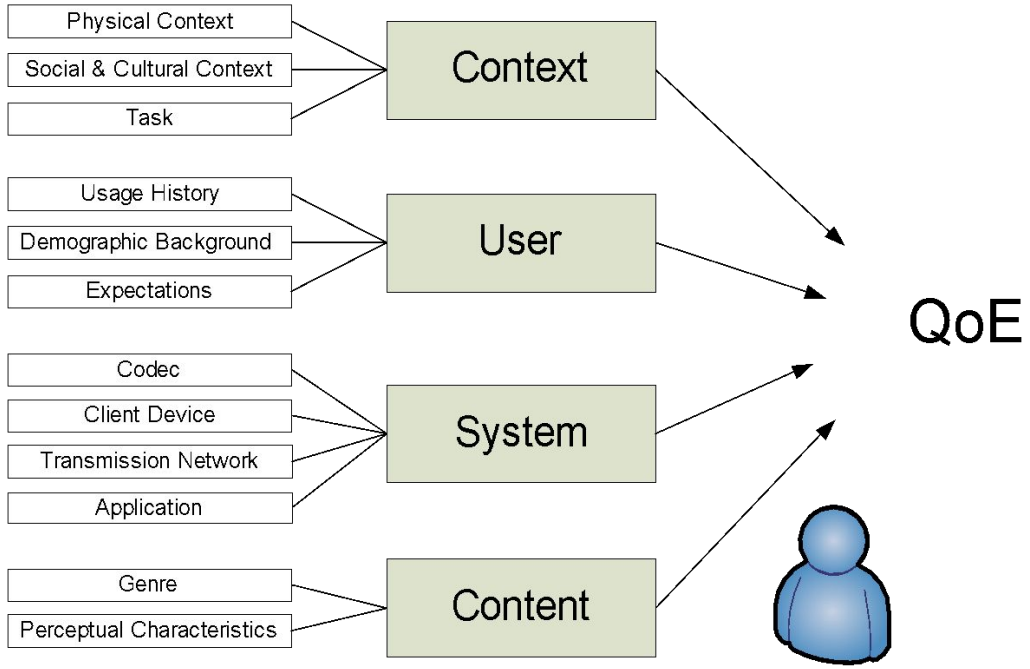


VQEG

Definitions: Quality of Experience

Quality of Experience (QoE) is the degree of delight or annoyance of the user of an application or service.

- It results from the fulfillment of expectations with respect to
  - the utility and / or
  - enjoyment of the application or service
  - in the light of the user's personality and current state.
- QoE focuses on the entire service experience
- Holistic concept, with its roots in telecommunication
- Proposed by Qualinet, current ITU-T Rec. P.10/G.100



VQEG

Definitions: User-Reported QoE

- Quality of Experience is an holistic concept addressing the user degree of delight or annoyance.
- In practice, we are only interested on the **effect of the manipulation of a few system conditions** (bitrate, network errors, etc.) in the user delight or annoyance.
  - Typically measured by user ratings in 1-5 scale
  - Can be used as the "quality of the data" (a user-related cost function for network optimization)

We define this restricted version of the QoE as **user-reported Quality of Experience**

We define **user-reported QoE** as
the quantification of the **impact of a system** on user delight or annoyance,
through **self-report**, behavioral, or psychophysiological studies.

This impact can be caused by the a**pplication, network, or hardware**
and is moderated by the **usage context.**

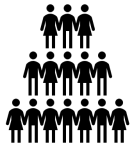
VQEG

Definitions: QoE metric and modeled QoE

A (user-reported) **QoE metric** is a quantitative measure that assesses the user-reported QoE statistically. Examples are the Mean Opinion Scores (MOS), the ratio of users rating good-or-better (GoB), the ratio of users rating poor-or-worse (PoW).

The **modeled QoE** is the output of a QoE model, which is based on or predicts a user-reported QoE metric. A **QoE model** considers various input signals and parameters to predict (user-reported) QoE in terms of a (user-reported) QoE metric. We assume that modeled QoE can be instrumentally measured using a QoE model, in the absence of subjective ratings.

Definitions: QoE hierarchy



System QoE is defined as the assessment of the modeled or user-reported QoE of the users of a particular service or system from a provider's perspective over a dedicated time frame. Typically, system QoE relies on modeled QoE through objectively measurable parameters and appropriate QoE models. The **expected system QoE** is a system QoE metric, quantifying the average QoE rating of an arbitrary user in the system.

Quality of Experience (QoE) is the degree of delight or annoyance of the user of an application or service. **User-reported QoE** is the quantification of the impact of a system on user delight or annoyance. **Modeled QoE** is the output of a QoE model, which is a perceived QoE metric.

Key Quality Indicators (KQIs) are metrics that directly or indirectly reflect the overall quality of an end-to-end service, which is related to a specific service or application.

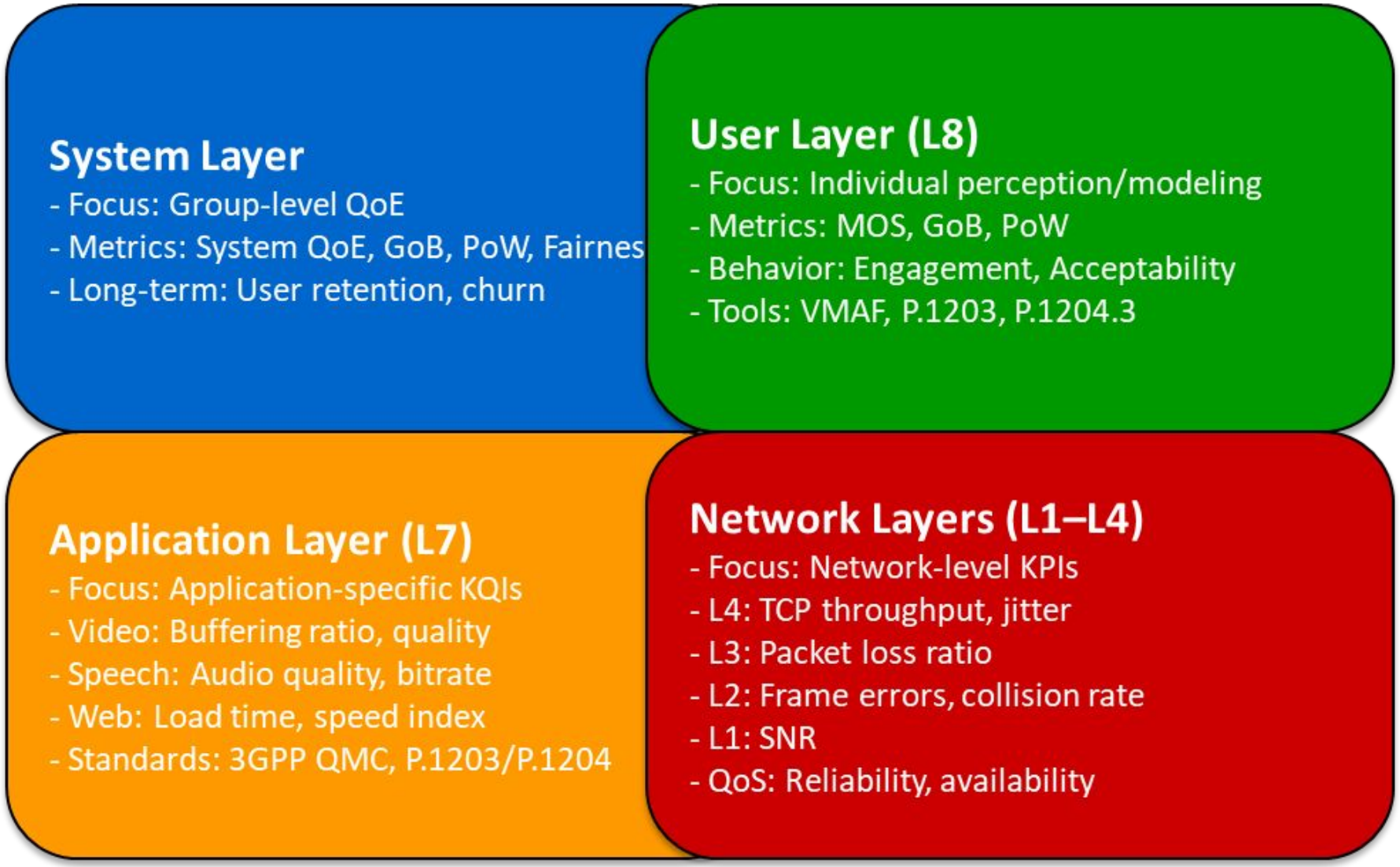
A **Key Performance Indicator (KPI)** is a specific type of network layer metric used to measure and evaluate the performance of a system or service on network level. KPIs are collected from the network or calculated from network measurements.

VQEG

Layered Approach

Layer	Information
System Layer	**System QoE:** QoE of a group of users of a particular service or system **System QoE Metrics**: Expected system QoE, System GoB, System PoW, QoE Fairness
Layer 8: User Layer	QoE scores based on - subjective ratings (user-reported **QoE**), e.g. MOS = average of opinion scores, e.g. GoB = Prob(„Good-or-better“) or - QoE models (**modeled QoE**) mapping KQIs and KPIs to modeled QoE metrics, e,g. MOS = function(KQIs, KPIs)
Layer 7: Application Layer	Key Quality Indicators (**KQIs**), for example, Video streaming: rebuffering ratio, video quality, reception ratio Web browsing: page load times, speed index
Layer 1 – 4: Network Layers	Key Performance Indicators (**KPIs**): measures on physical (L1), link (L2), network (L3), transport layer (L4) and potentially aggregated into KPIs L4: TCP goodput or throughput, and variability thereof (i.e. jitter) L3: IP packet loss ratio L2: Collision Rate, Frame Error Rate L1: Signal-to-Noise Ratio (SNR)

Layers


System Layer
- Focus: Group-level QoE
- Metrics: System QoE, GoB, PoW, Fairnes
- Long-term: User retention, churn
User Layer (L8)
- Focus: Individual perception/modeling
- Metrics: MOS, GoB, PoW
- Behavior: Engagement, Acceptability
- Tools: VMAF, P.1203, P.1204.3
Application Layer (L7)
- Focus: Application-specific KQIs
- Video: Buffering ratio, quality
- Speech: Audio quality, bitrate
- Web: Load time, speed index
- Standards: 3GPP QMC, P.1203/P.1204
Network Layers (L1–L4)
- Focus: Network-level KPIs
- L4: TCP throughput, jitter
- L3: Packet loss ratio
- L2: Frame errors, collision rate
- L1: SNR
- QoS: Reliability, availability

VQEG

QoS to QoE Models

standards for real-time applications

Video

video-based services, including video, audio quality estimation and quality integration (ITU-T P.1203, P.1204, 3GPP TR 26.909…)

Audio

VoIP audio performance (ITU-T G.107, P.1305, P.1310…)

Gaming

Cloud and terminal-based games, under error-prone low-latency networks (ITU-T G.1072, EEE P2948/P2949…)

Telemetry & QoS/QoE planning

Proactive analysis of network performance and support for customer service troubleshooting (ITU-T Y.1541, GSTR-5G QoE, J.1631…)

Metaverse AR/VR/XR

Perceived experience of virtual reality (VR) and augmented reality (AR) services (ITU-T Y.3109, P.1320, IEEE P2048…)

VQEG

Considerations

Approaches	Considerations
Technical context	How the model incorporates the **technical context** (screen size, codec, devices OS, N/W impairments, etc.) and the user context (e.g. how different users may rate the same technical conditions)
Accuracy	What is being measured, the required level of accuracy when using a model (what is accounted for and not)
Limitation of the models	Understand the limitations of the model (may not account for user preference, vision acuity, artistic content , user mood - right content vs video quality)
Absolute vs Relative	Absolute score vs relative (reference to a known source). Some models are absolute (e.g. No-Reference metrics) and others are relative (typical Full-Reference metrics). User ratings, even in absolute scale, may be relative to rating context (e.g. the set of qualities which appear in the rating experiment) and expectations may change with time (e.g. what is “excellent” quality today may be rated as “good” next year, if a better technology appears).
Impairments	Which kinds of impairments / System-related factors should be covered by the model? Is the model accounts for n/w impairments (e.g. short form videos: stalls or re-buffering;video calls: latency/jitter)
Context factors	Which are the right context factors to take into account (mobile vs fixed/home,business vs entertainment)

Discussion

- Questions or comments

- In what way could QoE be useful?
  - We can only understand it partly through user-reported or modelled QoE. Do we need something more?

- QoE works on an average level, but what about personal preferences and requirements?

VQEG

Industry Alignment

CAPs <> CSPs

VQEG

The Need for QoE Management

- Core Principle: QoE is intrinsically End-to-End.
- Challenge: No single entity has full E2E visibility or control.
  - CAPs control: Application, Content processing/delivery (partially), Client software.
  - CSPs control: Network infrastructure (access, core), Interconnects.
  - User controls: Local environment (device, LAN/WiFi).

How to optimize QoE without full E2E control?

Collaboration & Info Exchange

VQEG

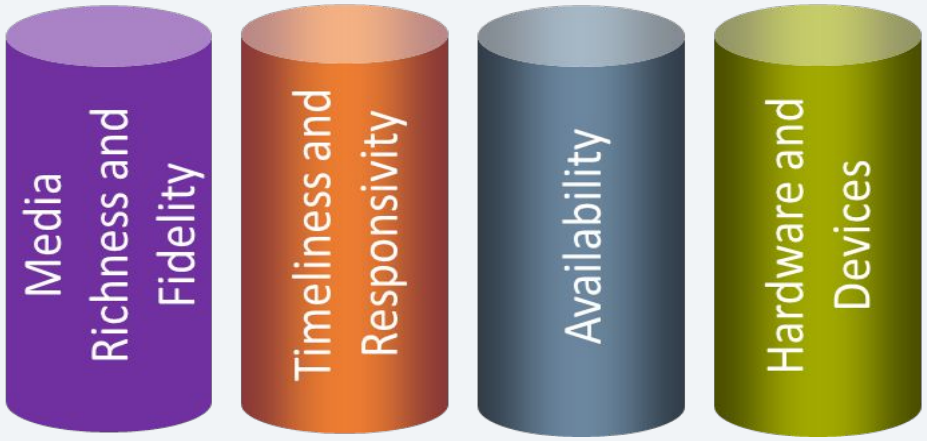
QoE Influencing Factors and Category

QoE

Human Related

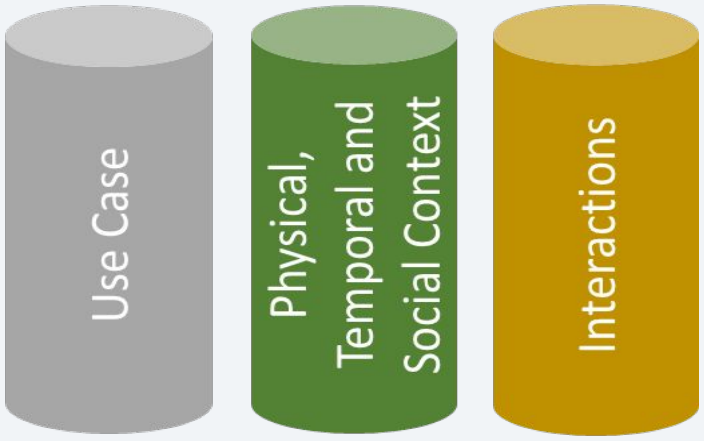
- Vision acuity
- Hearing acuity
- Tactile acuity
- Cognitive processing abilities
- Body posture and movement
- Age, gender,
- Tendency to experience immersion
- Emotional state, stress, attitude
- Susceptibility to simulator sickness (motion mis-match delay, vestibular system and visual system)
- Realism/style
- Video background plausibility
- Locomotion (movement)
- Proxemics (interpersonal distances)

System Related



- Audio/video fidelity
  - Blur/blockiness
  - Freeze frames/stalling
  - Speech distortion, echo, loudness
  - HMD Spatial resolution
  - Field of View
- Responsiveness / Latency
  - MTP (Motion-to-Photon)
  - A/V synchronization delay
  - Initial start/buffering
  - Response time of interactive operation

Context Related

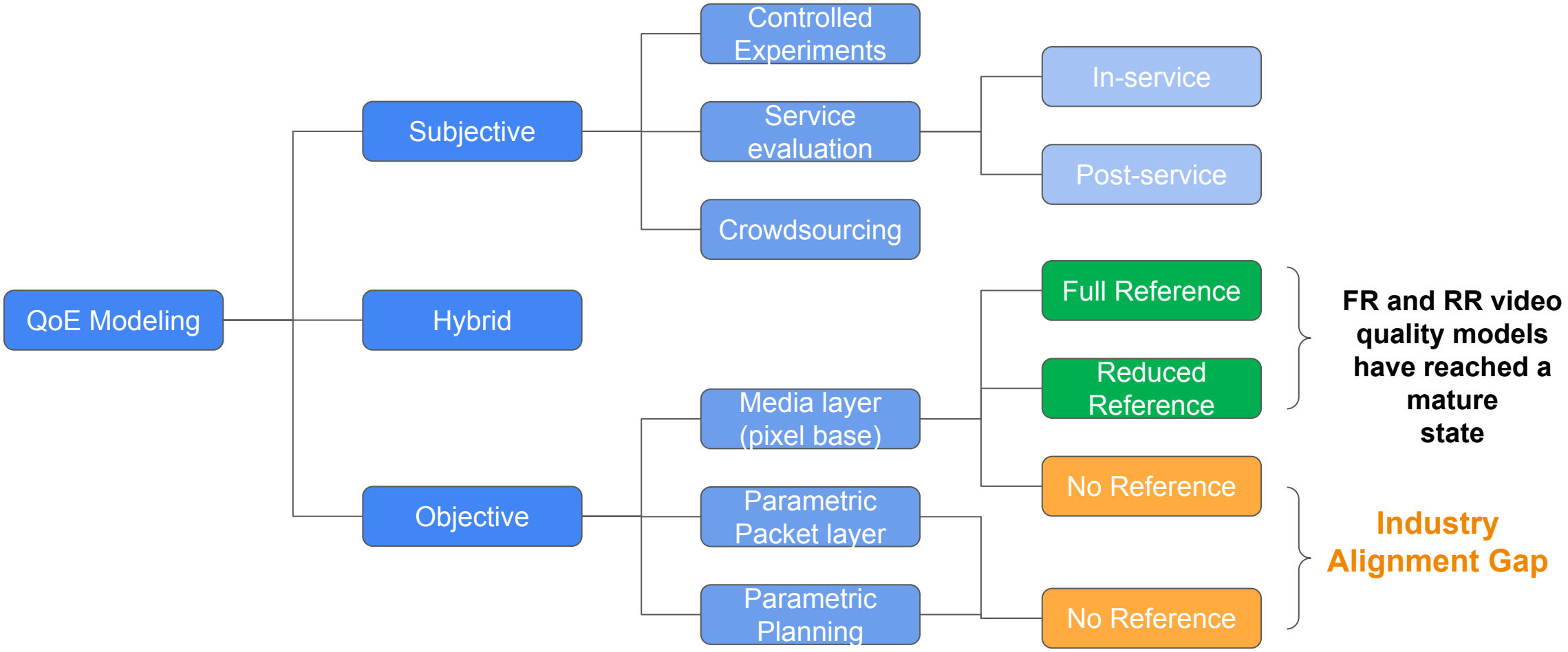


- Interactivity level
- Task type: Professional, social, leisure, entertainment
- Collaboration level
- Sharing content-space
- Objects in space
- Degree of realism
- Geometric/occlusion , illumination consistency
- Eye/head tracking
- Content quality , artistic quality and interest
- Frequency and duration of use
- How service is consumed : individual/ group
- Wired/wireless
- Room lighting - location
- Device type
- Hand gesture recognition accuracy and delay

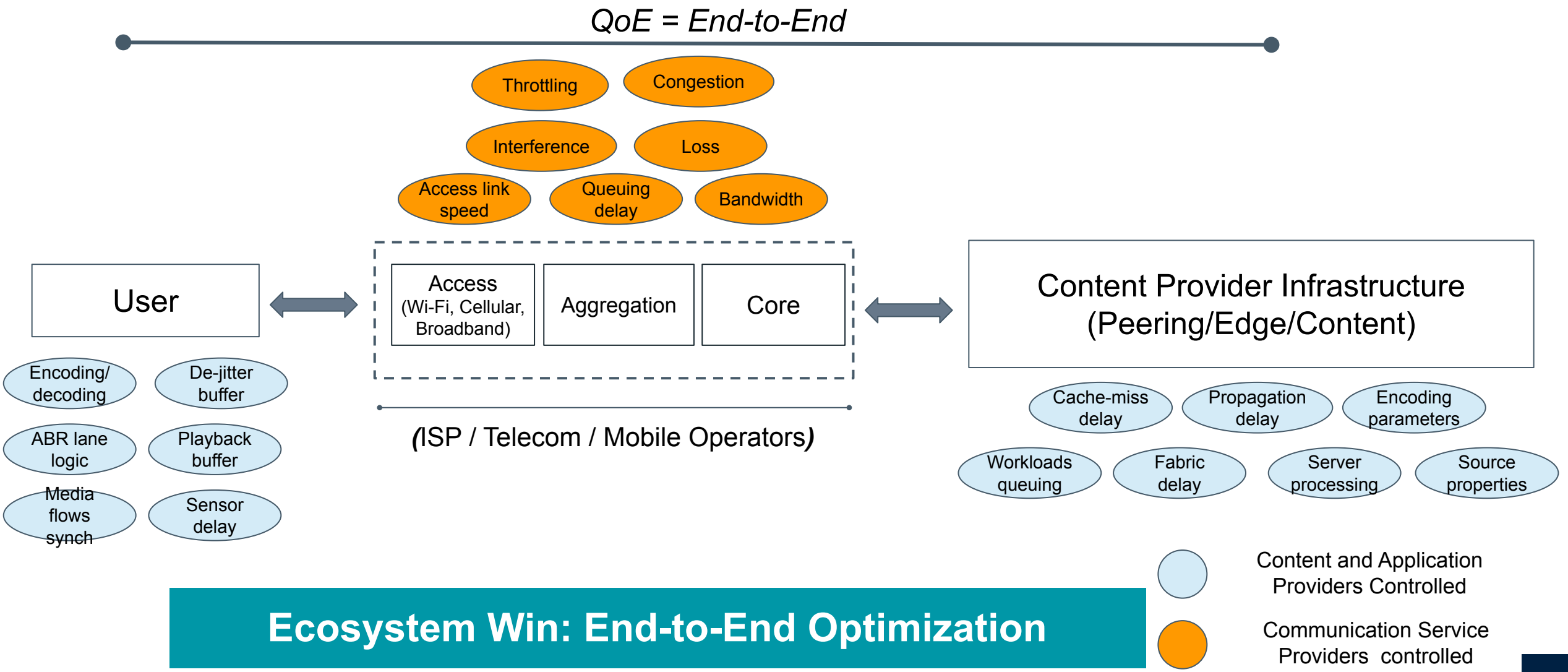
Industry Standardized QoE Models/Metrics

Application	QoE Metrics Examples
Conversational VoIP Call	**Voice QoE = f (Responsivity, Media Fidelity)** One way delay speech distortion, Synchro echo and sound level **Example of well accepted QoE metrics/measurement methods: MOS, PESQ (ITU-T P.862), POLQA(ITU-T P.863), E-model R-factor (ITU-T G.107)**
Video	**Video QoE = f (Temporal video quality, Spatial video quality, Timeliness / Responsivity Context)** fluidness/jerkiness media fidelity, blur/blockiness, initial loading, Stalls (#, duration, timing), one-way delay, A/V Synch (conversational) content type **Example of well accepted QoE metrics/measurement methods: PSNR, SSIM, VMAF, ITU-T G.1070/71, ITU-P.1204**
Gaming	**Gaming QoE = f (Temporal video quality, Spatial video quality, Responsivity, Context, Human)** fluidness/jerkiness media fidelity, blur/blockiness, action motion response Stalls (#, duration, timing) game classification frame Loss Sensitivity Delay sensitivity **Example QoE metrics/measurement methods: R-factor for cloud gaming (ITU-T G.1072/G.1032)**
3D VR/SR/XR	**3D Virtual Reality QoE = f (Temporal video quality, Spatial video quality, Responsivity, Context Human)** fluidness/jerkiness media fidelity, blur/blockiness, motion- to-photon Stalls (#, duration, timing) content interest, task type: collab or individual Locomotion immersion, motion sensitivity degree of realism **Emerging/under development QoE metrics/measurement** **2023 ITU-T PSTR-OQMXR "Objective quality modelling for XR services"**

Video QoE Model Approaches



Factors that are impacting QoE Management

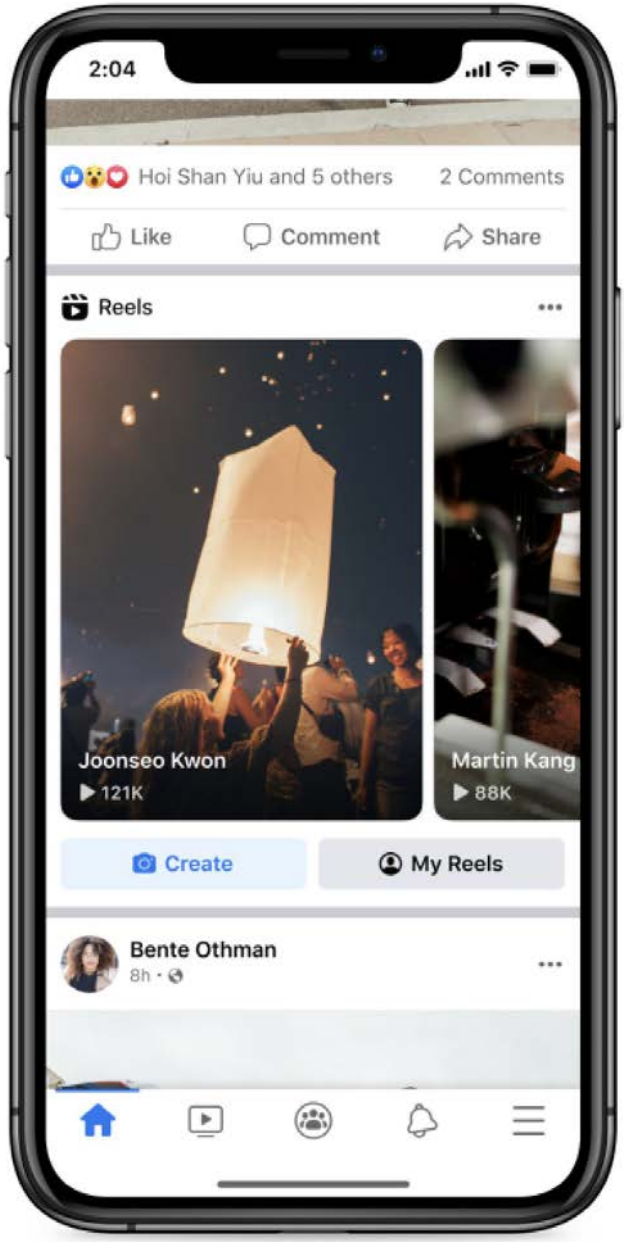


Ecosystem Win: End-to-End Optimization

VQEG

SFV (Short Form Video) Key QoE Aspects

Timeliness	• Initial loading, startup delay, stalls, rebuffering
Temporal Quality	• Fluidness, jerkiness [1] • AV Synch quality
Audio Quality	• Media Fidelity (Speech distortion, synch, frequency range, etc)
Spatial Quality	• Pixel fidelity • Blurriness, blockiness
Context	• Content type (static images, dynamic video) • Access network (cellular, wi-fi, home, moving)
Measurable QoE Metric	• FR Industry standard QoE metrics for video fidelity: PSNR, SSIM, VMAF • Business metrics: Retention ratio, User Watch Time • Human Perception: User evaluation



[1] (not as critical with reliable transport protocol with retransmission, more attributed to client (HW, buffering, battery)

QoE Aware Management

Visibility: visibility of what is happening and what could be happening if changes are implemented,

- CSPs: monitor the health of the network and predict QoE impact of network configuration
- CAPs: balance optimal codec bit rate for a given network condition condition

Efficiency: Optimal balance between delivery quality and resource use.

- CSP: Minimize resource usage (spectrum, equipment, cost) without waste.
- CAP: Efficient Content Data Network, compute, encoding resource usage; data egress.

Fairness: Moving beyond simple Bandwidth Fairness to QoE Fairness.

- Problem: Equal bandwidth ≠ Equal QoE (different app needs)
- Goal: Allocate resources to achieve fair experience across users/services.

Taking actions based on visibility, efficiency and fairness goals

Discussion

QoE is End-to-End and requires multi-dimensional optimization

Bridging the gap between CAPs and CSPs to achieve a win-win, any other thought?

How to improve video service delivery: Visibility, Efficiency and Fairness

QoE Management Framework

Proposed CSP/CAP Metrics, Models & Tools

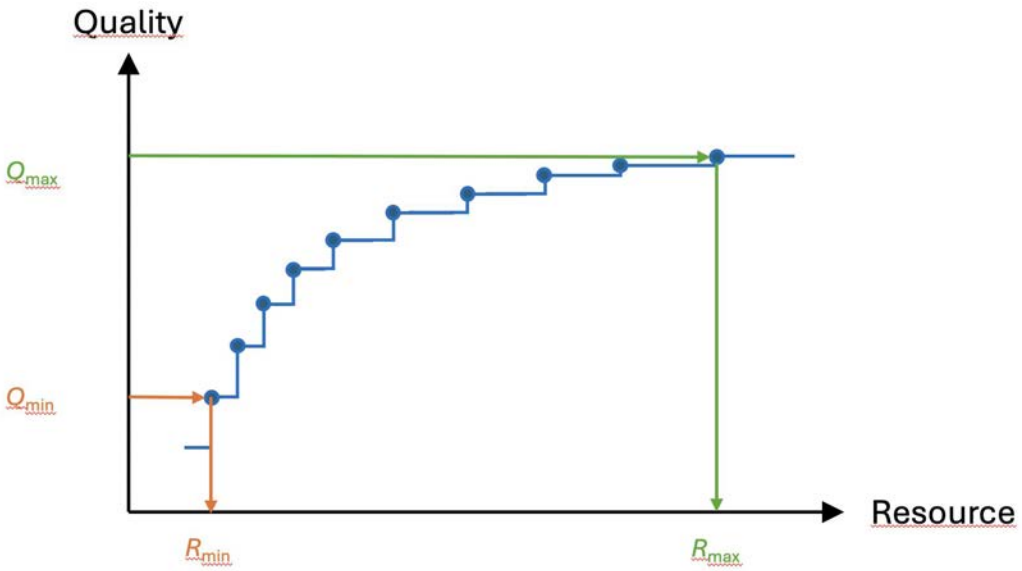
VQEG

Goals and Requirements

Goal: Define concrete metrics, models, and tools for CAP-CSP exchange to improve quality on a *Resource-Quality curve*

Requirements:

- **Common & Understood:** Agreed metrics, well-known protocols, standardized and non-proprietary.
- **Standardized Scale:** Consistent numerical representation (e.g., 1-5 MOS, 0-100 score).
- **Subjective Correlation:** Proven link to user perception (low bias, documented accuracy - RMSE, Pearson).
- **Implementation:** Calculable by CSP (with CAP data) or provided directly by CAP (via defined interface, e.g., CMCD extension possibility).



→ Understand relationship (e.g., Bitrate vs. Quality).

→ Identify key points:

- Qmin/Rmin (minimum acceptable)
- Qmax/Rmax (diminishing returns/avoid waste or overprovisioning)

VQEG

General Components & Shared Metrics

Common QoE components across services:

- Media Fidelity (visual/audio quality)
- Media Delivery Continuity (stalls, freezes, drops)
- Media Delay & Interactivity (latency, responsiveness, startup)

CAP Provided

- Current Quality (Video/Audio/Stall scores)
- Cost Function (Rate-Distortion curve, Qmin/Rmin, Qmax/Rmax)
- Metadata (codec, bitrate, ...).

CSP Provided

- Network State (Congestion level, ECN)
- Subscriber Entitlement (Policy Rate Limits)

VQEG

Proposed Framework – Shared State Table

Scope Limitations:

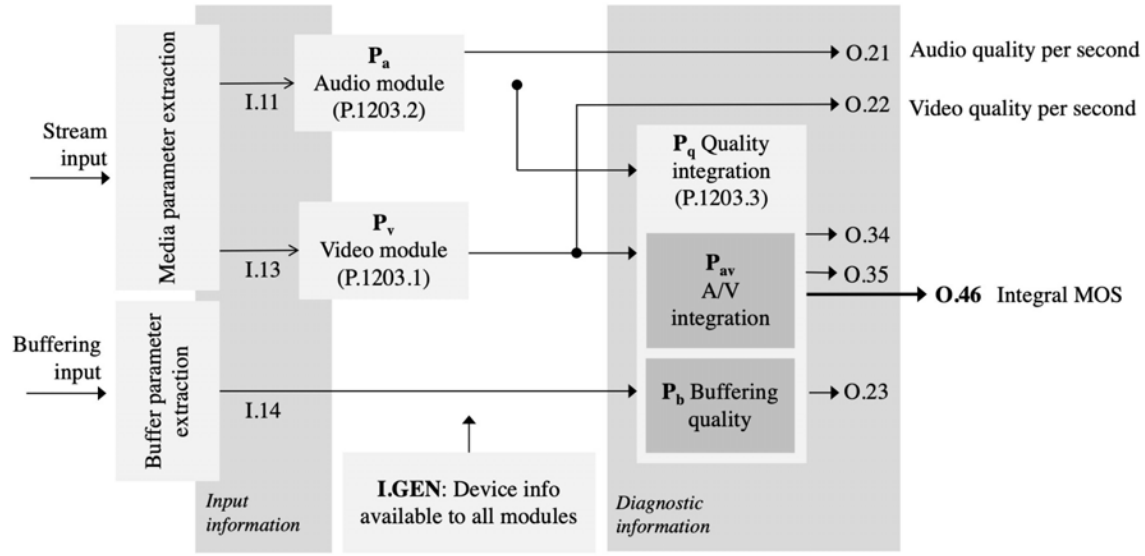
- Video (short/long form)
- Conferencing/interactive services

Core Idea: Logical shared view of QoS/QoE status between CAP & CSP.

Mechanism:

- Maintains continuous state (potentially divergent views of same params like BW).
- One-way updates on state changes (more efficient than request/response).
- Implementation potential: Side-channel, metadata exchange (e.g., via NWDAF).

Leverage existing models for semantics: e.g., use ITU-T P.1203 structure as a template for needed information types like “per-second video quality”.



Use case: Long Form Video Services (VoD)

Layer	Media Fidelity	Media Continuity	Media Interactivity	Provider
User (8)	Integral MOS (P.1203)			CAP
	Video Fidelity (VMAF) Audio Fidelity (PEAQ)	Stall duration (ms)		
Application (7)	Bitrate ladder			
	Media metadata (codec, bitrate, resolution, frame rate)			
Transport (4)	Rsus: Maximum sustainable throughput in the network			CSP
	Subscriber network entitlement: Video Policy Rate Limit, Entitlements TS.43			
Network (1-4)	Network congestion: % Resource utilization (radio, IP…), ECN, Congestion Flag, buffer queue buildup			

VQEG

Use case: Short Form Video Services

Layer	Media Fidelity	Media Continuity	Media Interactivity	Provider
User (8)	*Integral MOS (P.1203)*			CAP
	Video Fidelity (VMAF) Audio Fidelity (PEAQ)	Stall duration (ms)	Start-up Delay	
Application (7)	Bitrate ladder			
	Media metadata (codec, bitrate, resolution, frame rate)			
Transport (4)	Rsus: Maximum sustainable throughput in the network			CSP
	Subscriber network entitlement: Video Policy Rate Limit, Entitlements TS.43			
Network (1-4)	Network congestion: % Resource utilization (radio, IP…), ECN, Congestion Flag, buffer queue buildup			

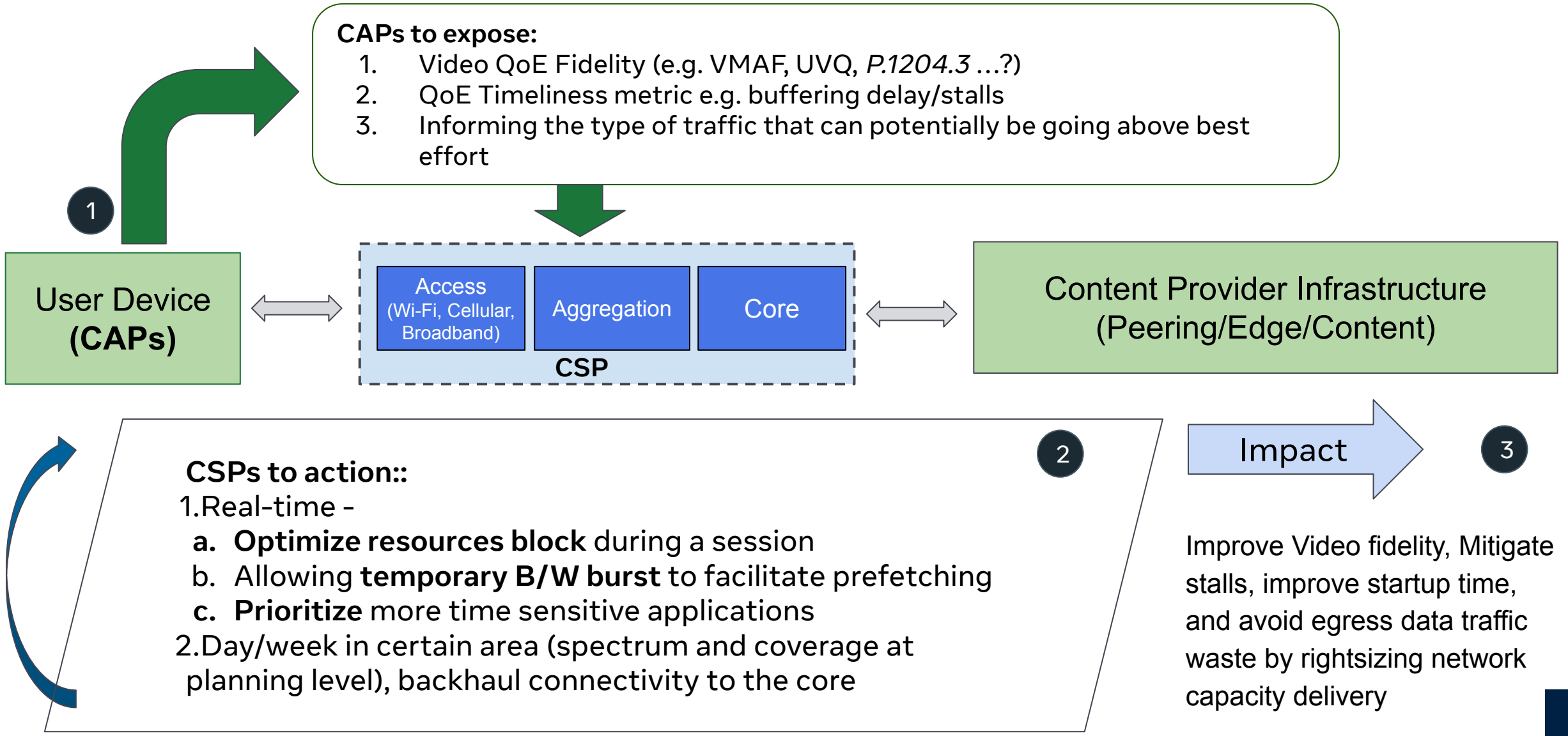
VQEG

Use case: Interactive Services (e.g. gaming)

Layer	Media Fidelity	Media Continuity	Media Interactivity	Provider
User (8)				CAP
	Video Fidelity (VMAF) Audio Fidelity (PEAQ)		Interaction Quality	
Application (7)	Bitrate ladder	Application KQIs (frame loss)	Interaction cost function	
	Media metadata (codec, bitrate, resolution, frame rate)			
Transport (4)	Rsus: Maximum sustainable throughput in the network			CSP
	Subscriber network entitlement: Video Policy Rate Limit, Entitlements TS.43			
Network (1-4)	Network congestion: % Resource utilization (radio, IP…), ECN, Congestion Flag, buffer queue buildup			
	Network statistics: Histograms for RAN packet delay DL/UL, … [3GPP TS 23.288]			

VQEG

Short Form Videos: CAPs → CSPs Actionable Control



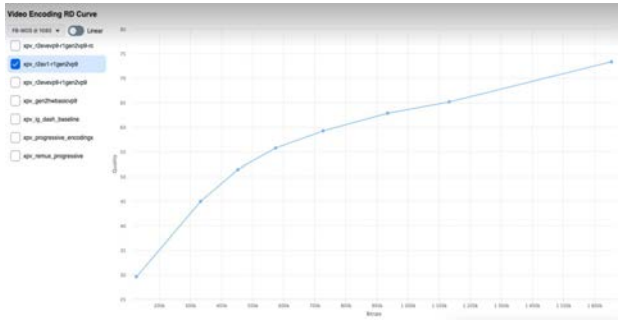
VQEG

Short Form Videos: CSPs → CAPs Actionable Control

2

CAPs to action::

1. select the optimal ABR track, configure ABR algorithm, or tune the delivery rate.
2. CAP can self-regulate the entire delivery.



Impact 3

Mitigate stalls, improve startup time, and avoid waste by modulating appropriate buffer size.

User Device **(CAPs)**

Access (Wi-Fi, Cellular, Broadband) | Aggregation | Core

CSP

Content Provider Infrastructure (Peering/Edge/Content)

1

CSPs to expose:

1. subscriber's network entitlements including video policy (data plan, e.g. 1.5-2.5mbps)
2. CSP can expose network state such as congestion level, localized events or flash crowd alerts.
3. Customer support

VQEG

Discussion

- Which are the benefits and risks of implementing an approach like this one?

- What are the main potential blockers?
  - Business alignment
  - Privacy
  - Complexity
  - Regulation
  - Standardization

- What could be the best mechanisms / protocols to implement a *shared table view* between CAPs and CSPs?

VQEG

Conclusions

and Next Steps

VQEG

Conclusions and Next steps

Conclusions:

- Clear need for **aligned,** practical QoE definitions & **metrics** for CAP/CSP **collaboration**.
- Proposed **framework** based on **shared state table** and semantic building blocks offers a viable path.
- **Information exchange** enables tangible benefits (efficiency, fairness, improved QoE).

Next Steps:

- Release the white paper (target summer 2025)
- Development of NEW QoE model to improve QoE management to fulfill gaps
- Validation: Proof of concept /lab test/Field trial /simulation on basic conditions
- Long term: VQEG/ITU-T SG12 Development of QoE-QoS model that is standardized in the CAPs/CSPs community

VQEG

Thank You

VQEG

Challenges and Gaps in Delivering Video QoE

1. Lack of commonly accepted QoE models, metrics and targets for short form videos
    a. Video quality metric only (consistent delivery without stalls, re-buffering)
    b. QoE timeliness metric (accounts for temporal artifacts during a viewing session: preloading delays, stalls
    c. There are no sufficiently consolidated metrics including both the fidelity aspect and timeliness.
    d. P.1203/1204 are existing standards that account for stalls (but not preloading delays)
2. Common metric definition among CAPs, among CSPs, between CAPs and CSPs
3. CAPs - CSPs Visibility: Metrics exchange to improve video delivery efficiency
    a. QoE results (e.g. MOS/VMAF) to be generated on short form videos
    b. QOS metrics (e.g.congestion level, policy treatment) to enable better QoE video delivery efficiency.
    c. Define what sort of action we can perform based upon the metrics we exchange
    d. Interface/API , protocol and format
4. Standardization
    a. There are no recommendations or standards that define testing methodologies for short form videos
    b. No Reference (NR) model - Still limited in their accuracy and not yet universally adopted
    c. Full reference : CSPs who do not have easily access to source content may be difficult to implement

CAPs Metrics and Actionable Control Insights
— Use Case: Short Form Videos

QoE Aspects	QoE metrics	How to measure	CSPs Actionable Control	CAPs Actionable Control
Audio quality	Audio Fidelity	POLQA, ITU P.1203, Codec type, Bit rate	subscriber's network entitlements including video policy (data plan, e.g. 1.5-2.5mbps) network state such as congestion level, localized events or flash crowd alerts. Customer support	Playback Video QoE Bandwidth estimators ABR lane bitrate selection Codec selection De-jitter /playback buffer adjustment Egress data traffic volume
Timeliness	Click to play time (CTPT)	Measured on the client, the interval between the time when a user click a video and the time when the video starts to play on the screen		
	Play success rate n (PSRn)	Percentage of SFV views which has a CTPT less than n seconds		
	Stalls	Measured on the client side per viewing session by some/combination of (1) number of stalls (longer than xxx ms) (2) total time of stalls (milliseconds) (3) meantime between stalls during a session		
Temporal quality	Fluidity	Measure on client by number of frames per second		
	Synchronicity	Measured on the client side per viewing session by some/combination of (1) numbers of audio/video out-of-synch (2) total time of audio/video out-of-synch (3) meantime between audio/video out-of-synch		
Spatial quality	Video fidelity	No Reference: Under Development Full Reference: PSNR, SSIM, or VMAF		
Context**	Client Device, Location	Display resolution and audio fidelity, mobile or stationary, network type		

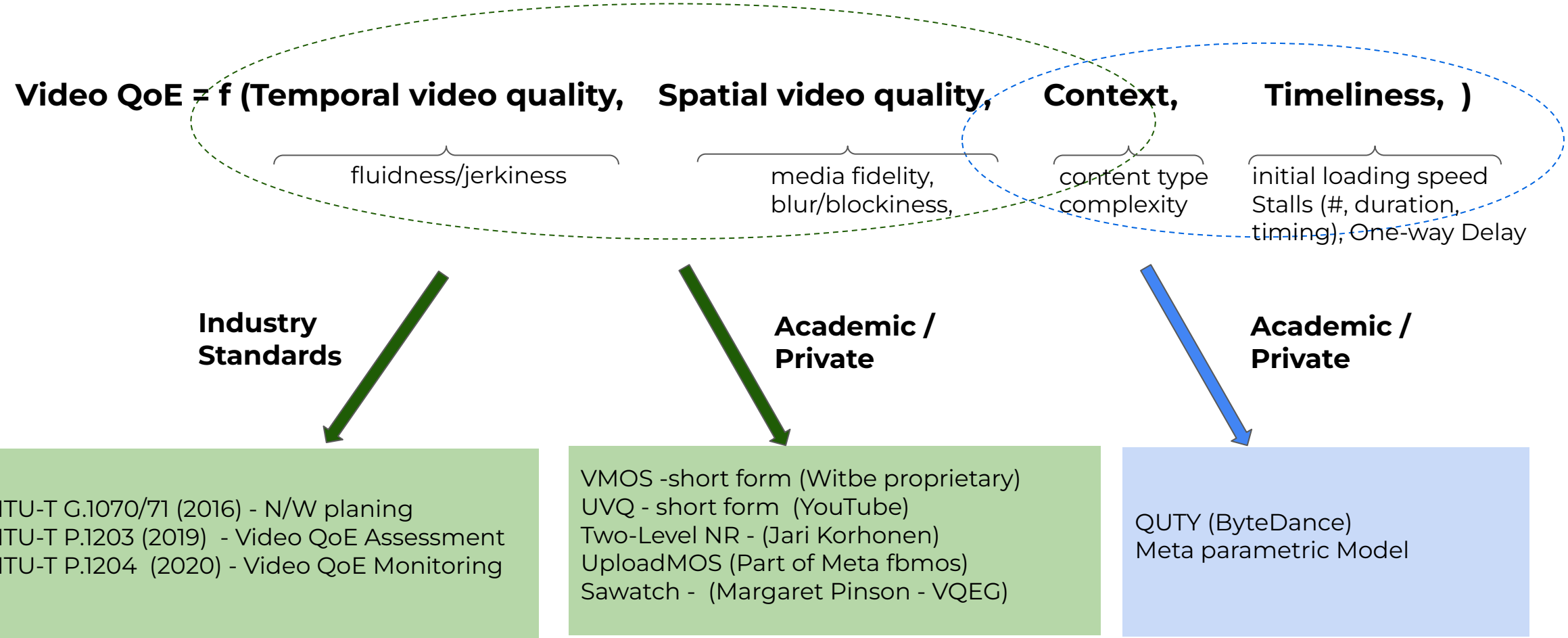
** The context may impact target values of QOE metrics for what an acceptable/good/excellent QoE. Metrics and tolerances can differ for WiFi (unconstrained BW) vs Cellular

CSPs Metrics and Actionable Control *(The CAPS View)*

— Use Case: Short Form Videos

QoS Aspects	QoS metrics	QoS Targets	CSPs Actionable Control
Bandwidth	Link speed/bitrate Session Bit Rate Policer/Shaper CIR, CBS, Queue depth Radio link signal (TDD/FDD, radio bandwidth)		Network loading (link utilization) Congestion level (PRB utilization, # of users) Traffic management policies Policy treatments per subscribers categorie QoS service class prioritization (e.g. QCI) Per geographic location (e.g. cell location)
Latency	RTT (Round Trip Time) One way delay Propagation delay (distance based)		
Queuing Delay	Buffer depth Link utilization Scheduling discipline Congestion level (PRB utilization, #of user per cell)		
Packet Loss	Buffer depth Policer/Shaper : CIR, CBS, Queue depth Link utilization Scheduling discipline Congestion level (PRB utilization, #of user per cell)		

NO Reference VIDEO QoE Models



QoE <> QoS Correlation

Exchanging Metrics to Gain End-to-End Ecosystem View

- **Use Case : Short Form Videos**

CAPs		
QoE Aspects	**QoE metrics**	**How to measure**
Audio quality	Audio Fidelity	POLQA, ITU P.1203, Codec type, **Media rate**
Timeliness	Click to play time (CTPT)	Measured on the client, the interval between the time when a user click a video and the time when the video starts to play on the screen
	Play success rate n (PSRn)	Percentage of SFV views which has a CTPT less than n seconds
	Stalls	Measured on the client side per viewing session
Temporal quality	Fluidity	Measure on client by number of frames per second
	Synchronicity	Measured on the client side audio/video out-of-synch per viewing session
Spatial quality	Video fidelity	No Reference: No industry standard Full Reference: PSNR, SSIM, or VMAF
Context	Client Device, Location	Display resolution and audio fidelity, mobile or stationary, network type

CSPs	
QoS Aspects	**QoS metrics**
Bandwidth	Link speed/bitrate Session Bit Rate or **Media Rate** Policer/Shaper CIR, CBS, Queue depth Radio link signal (TDD/FDD, radio bandwidth)
Latency	Round Trip Time One way delay Propagation delay (distance based)
Queuing Delay	Buffer depth Link utilization Scheduling discipline Congestion level (PRB utilization, #of user per cell)
Packet Loss	Buffer depth Policer/Shaper : CIR, CBS, Queue depth Link utilization Scheduling discipline Congestion level (PRB utilization, #of user per cell)

How can CSPs and CAPs exchange QoE-QoS metrics

1. CAPS agreed on a common metric to be shared
2. An easy-to-understand metric that captures the QoE for a user/viewer of a given service, at the session level (or some other unit of time), at a common numerical scale ([0,100] ?) - on a mobile device
3. Metric needs to be correlated with subjective opinions, with zero-bias and a well documented accuracy (standard deviation)
4. Metric needs to be either easily calculated by the CSP independently or provided by each CAP via a commonly agreed upon interface (in-band or out-of-band)
5. QoE to QoS correlation needs to be well defined
6. Offer multiple implementations based on existing video quality metrics - ideally, open-source (SSIM, VMAF, FUNQUE, UVQ, Others?)

Metric needs to be actionable, optimization tradeoffs need to be validated and understood

A Way Forward - How do we get organized as as group?

1. Agree on business motivation (More efficient End-to-End video delivery?, improve QoE?)
2. Agree on scope (generic vs bite size problem)
    a. improve apps QoE on all Network types?)
    b. Start with a common metric among CAPS, same for CSPs
3. Identity gap(s) that VQEG can contribute and influence industry
    a. Establishment of a framework to define QoE/QoS metrics, targets and measurements
    b. Common QoE-QoS testing methodology
    c. Metrics to be exchanged among CSPs, CAPS and between
    d. Industry alignment on above
4. Publish a white paper/journal publication
    a. To address gap above
    b. that explains the details of the QoE-QoS Framework, metrics and its implementations, as well as correlation with subjective studies
5. VQEG to contribute to the effort of Standardization through an existing SSO or liason:
    a. E.g. Conduct testing, development of QoE mode)
    b. ITU-T (SG12), AOM (Alliance for Open Media), IETF, Others ?
6. Create a working group to address and prioritize gap(s)

BACKUP

Workshop Topics - This slide will be removed

Views from Communication Service Providers (CSPs) and Content Application Providers (CAPs)

Industry alignment - what, why

Vision What we would like to see from the content providers or N/W operators - what is our vision
North Star: Develop QoE-QoS models And/OR develop framework to manage QoE in 5G/6G

Use cases
1- Streaming (Netflix),
2- interactive (cloud gaming, short form video)
3- video telepresence (conversational - more emphasis on Audio?)
Requirements from both QoE and QoS perspective
Information exchange based upon the type of traffic, Targets, provide context on the type of traffic

What are we willing to contribute to improve in general QoE
QoE and QoS Metrics and Influencing factors
Metrics to be exchanged Eg 3GPP metrics FMK
Metrics and targets
Ex: CSPs provide QoS level
Granularity, how often do we probe (often we don't have small time windows, Per min? Per 15min? Per hour, Per session for SFV
Limitation because of Hw limitation

What sort of action we perform based upon the metrics we exchange
How to interpret QoE/QoS metrics and how **actionable for operators and content providers**
Establish FMK, trade offs , e.g. can decrease delay for all apps
- Map QoE to QoS, what are the most important influencing factors
- Should i increase the BW, decrease latency, which control knobs
- Parametric models ITU p.1204
- provide hint (ex my QoE is bad , i think it is because of a,b,c)

Review existing standards 3GPP