



Technical Report

TR-499

Broadband Service Metrics

Issue: 1

Issue Date: October 2024

Notice

The Broadband Forum is a non-profit corporation organized to create guidelines for broadband network system development and deployment. This Technical Report has been approved by members of the Forum. This Technical Report is subject to change. This Technical Report is owned and copyrighted by the Broadband Forum, and all rights are reserved. Portions of this Technical Report may be owned and/or copyrighted by Broadband Forum members.

Intellectual Property

Recipients of this Technical Report are requested to submit, with their comments, notification of any relevant patent claims or other intellectual property rights of which they may be aware that might be infringed by any implementation of this Technical Report, or use of any software code normatively referenced in this Technical Report, and to provide supporting documentation.

Terms of Use

1. License

Broadband Forum hereby grants you the right, without charge, on a perpetual, non-exclusive and worldwide basis, to utilize the Technical Report for the purpose of developing, making, having made, using, marketing, importing, offering to sell or license, and selling or licensing, and to otherwise distribute, products complying with the Technical Report, in all cases subject to the conditions set forth in this notice and any relevant patent and other intellectual property rights of third parties (which may include members of Broadband Forum). This license grant does not include the right to sublicense, modify or create derivative works based upon the Technical Report except to the extent this Technical Report includes text implementable in computer code, in which case your right under this License to create and modify derivative works is limited to modifying and creating derivative works of such code. For the avoidance of doubt, except as qualified by the preceding sentence, products implementing this Technical Report are not deemed to be derivative works of the Technical Report.

2. NO WARRANTIES

THIS TECHNICAL REPORT IS BEING OFFERED WITHOUT ANY WARRANTY WHATSOEVER, AND IN PARTICULAR, ANY WARRANTY OF NONINFRINGEMENT AND ANY IMPLIED WARRANTIES ARE EXPRESSLY DISCLAIMED. ANY USE OF THIS TECHNICAL REPORT SHALL BE MADE ENTIRELY AT THE USER'S OR IMPLEMENTER'S OWN RISK, AND NEITHER THE BROADBAND FORUM, NOR ANY OF ITS MEMBERS OR SUBMITTERS, SHALL HAVE ANY LIABILITY WHATSOEVER TO ANY USER, IMPLEMENTER, OR THIRD PARTY FOR ANY DAMAGES OF ANY NATURE WHATSOEVER, DIRECTLY OR INDIRECTLY, ARISING FROM THE USE OF THIS TECHNICAL REPORT, INCLUDING BUT NOT LIMITED TO, ANY CONSEQUENTIAL, SPECIAL, PUNITIVE, INCIDENTAL, AND INDIRECT DAMAGES.

3. THIRD PARTY RIGHTS

Without limiting the generality of Section 2 above, BROADBAND FORUM ASSUMES NO RESPONSIBILITY TO COMPILE, CONFIRM, UPDATE OR MAKE PUBLIC ANY THIRD PARTY ASSERTIONS OF PATENT OR OTHER INTELLECTUAL PROPERTY RIGHTS THAT MIGHT NOW OR IN THE FUTURE BE INFRINGED BY AN IMPLEMENTATION OF THE TECHNICAL REPORT IN ITS CURRENT, OR IN ANY FUTURE FORM. IF ANY SUCH RIGHTS ARE DESCRIBED ON THE TECHNICAL REPORT, BROADBAND FORUM TAKES NO POSITION AS TO THE VALIDITY OR INVALIDITY OF SUCH ASSERTIONS, OR THAT ALL SUCH ASSERTIONS THAT HAVE OR MAY BE MADE ARE SO LISTED.

All copies of this Technical Report (or any portion hereof) must include the notices, legends, and other provisions set forth on this page.

Issue History

Issue Number	Issue Date	Issue Editor	Changes
1	October, 2024	Ken Kerpez, DZS	Original

Comments or questions about this Broadband Forum Technical Report should be directed to info@broadband-forum.org.

Editor: Ken Kerpez, DZS

Work Area Director(s): Jonathan Newton, Vodafone

Project Stream Leader(s): Fabio Giudici, OutSys

Table of Contents

1. Executive Summary	6
2. Purpose and Scope	7
2.1 Purpose	7
2.2 Scope	7
3. References and Terminology	8
3.1 References	8
3.2 Definitions	11
3.3 Abbreviations	12
4. Technical Report Impact	14
4.1 Energy Efficiency	14
4.2 Security	14
4.3 Privacy	14
5. Introduction	15
6. Overall Service Metrics	15
7. Overall Service Metric Dimensions	18
8. Example Services and their Metrics	19
8.1 Service Metrics example: XR Rendering	19
8.1.1 <i>Applicable Service Metric Dimensions</i>	22
8.2 Service Metrics example: Multiple IoT Sensors	22
8.2.1 <i>Applicable Service Metric Dimensions</i>	24
8.3 Service Metrics example: High-Speed Internet (HSI) access	24
8.3.1 <i>Applicable Service Metric Dimensions</i>	25
8.4 Service Metrics example: Entertainment video	25
8.4.1 <i>Applicable Service Metric Dimensions</i>	27
8.5 Video Teleconferencing	28
8.5.1 <i>Applicable Service Metric Dimensions</i>	30
8.6 Ultra Reliable Low Latency Communications (URLLC)	30
8.6.1 <i>Applicable Service Metric Dimensions</i>	32
8.7 Network support for Speed Tests	32
8.7.1 <i>Applicable Service Metric Dimensions</i>	32
8.8 Service Metrics example: AI as a Service (AlaaS)	32
8.8.1 <i>Applicable Service Metric Dimensions</i>	34
8.9 Peer-to-Peer file sharing	34
8.10 Security cameras	35
8.11 Gaming	35
8.12 Web Browsing	35
Appendix I. Summary of existing standards to apply to service metrics	36
I.1 BBF TR-304	36
I.1.1 <i>TR-304 Metric Categories</i>	38
I.2 BBF TR-126	38
I.2.1 <i>TR-126 Metric Categories</i>	38
I.3 BBF TR-452.1, Quality Attenuation Measurement Architecture and Requirements	39
I.4 BBF TR-421, Application-Layer Test Traffic Architecture and Requirements	39

I.5	BBF-TR-471, Maximum IP-Layer Capacity Metric, Related Metrics, and Measurements.....	39
I.6	3GPP TS 28.554 KPIs.....	39
I.6.1	3GPP TS 28.554 Metric Categories.....	41
I.7	ATIS IPTV Interoperability Forum (IIF).....	42
I.7.1	ATIS-0800008.v002, Quality of Service Metrics for Linear (Broadcast) IPTV.....	42
I.7.2	ATIS-0800041, Implementers Guide to IPTV QoS Metrics	43
I.7.3	ATIS IIF Metric Categories.....	44
I.8	ETSI GS NFV-TST 008	44
I.9	5G-PPP White Paper: Beyond 5G/6G KPIs and Target Values	45
I.10	Cloud Service Metrics	45
I.11	IETF IPPM.....	46
I.12	IANA Performance Metrics Registry	46
I.13	Metro Ethernet.....	46
I.14	IEEE Std 802.3.2, IEEE Standard for Ethernet YANG Data Model Definitions	47

Table of Figures

Figure 1. Metrics, services, use cases, and dimensions.....	15
Figure 2 Quality layers for XR Rendering.....	20
Figure 3 Quality layers for Multiple IoT sensors.....	22
Figure 4 Quality layers for entertainment video.....	26
Figure 5 Quality layers for video teleconferencing.....	28
Figure 6 Quality layers for AlaaS.....	33

Table of Tables

Table 1. Overall Service Metrics.....	16
Table 2. Overall Service Metric Dimensions.....	19
Table 3. [TR-304 Table 2] – Static Access Service Attributes.....	36
Table 4. [TR-304 Table 4] – Static Access Service Attributes per protocol layer.....	36
Table 5. [TR-304 Table 5] – Dynamic Access Service Attributes.....	37
Table 6. [ATIS-0800041] Quality and Impairment scales, Five Grade Scale.....	43
Table 7. [ATIS-0800041] Alarm Levels.....	43
Table 8. [ATIS-0800041] Error Characterization	43

1. Executive Summary

Computing and networking are now being combined and elevated to function as services delivery platforms. Operators need to fully grasp solutions for assuring, monitoring, and troubleshooting emerging services, including delivery across wireline and 5G/6G access networks, content origination, computing, and sessions. Guidance is needed to manage service-level broadband delivery, with commonality across the industry to streamline deployment and combine service management with network management for service providers, network operators and end users.

This Technical Report identifies use cases for assuring broadband services delivery, including operations entities, technologies, service segments, and service types. Existing standards for metrics are identified and summarized. The application of existing standards for service assurance metrics are categorized and metrics for assuring services delivery are identified for representative services, and overall across all services. Service metrics are categorized into quality layers: Transaction quality, Content quality, Flow quality, and Network quality. Related dimensions that accompany metrics, such as various statistics and measurement methods are also identified.

This Technical Report provides wide-ranging sets of metrics that will assist in assuring services origination and delivery through the network.

2. Purpose and Scope

2.1 Purpose

This Technical Report addresses the needs of an operator to monitor and assure a platform for service delivery, encompassing QoS flows or slices across the network as well as vital aspects of service origination such as computing platforms, resources, and sessions.

2.2 Scope

This Technical Report identifies sets of service metrics or Key Performance Indicators (KPIs) that impact service delivery. This Working Text also describes use cases, services, and existing metrics that are used to identify service metrics. A categorization of service metrics is defined, and service metrics are identified. A metric may be described along many dimensions such as statistics used to represent each metric, measurement point, measurement time, etc. Service metric dimensions are identified here.

Service metrics go beyond network metrics and include session management, session failure rates, resource assignment, reliability, retainability, and utilization including network elements, network links, service origination, cloud, management and control, content quality, etc. These can support each service, QoS flow or service class, and include important metrics to represent the health and function of many phases of providing a service. Faults, performance, resource exhaustion and many other aspects of service delivery health can be identified with these metrics.

3. References and Terminology

3.1 References

The following references are of relevance to this. At the time of publication, the editions indicated were valid. All references are subject to revision; users of this Technical Report are therefore encouraged to investigate the possibility of applying the most recent edition of the references listed below.

A list of currently valid Broadband Forum Technical Reports is published at www.broadband-forum.org.

Document	Title	Source	Year
[1] TR-126	Triple-Play Services Quality of Experience (QoE) Requirements	BBF	2006
[2] TR-143	Enabling Network Throughput Performance Tests and Statistical Monitoring	BBF	2023
[3] TR-181	TR-181: Device Data Model	BBF	2023
[4] TR-304	Broadband Access Service Attributes and Performance Metrics	BBF	2015
[5] TR-421	Application-Layer Test Traffic Architecture and Requirements	BBF	2019
[6] TR-452.1	Quality Attenuation Measurement Architecture and Requirements	BBF	2020
[7] TR-452.2	Quality Attenuation Measurements Using L2 PM OAM	BBF	2022
[8] TR-452.5	Quality Attenuation Measurement Architecture and Requirements	BBF	2023
[9] TR-471	Maximum IP-Layer Capacity Metric, Related Metrics, and Measurements	BBF	2021
[10] TS 29.520	Network Data Analytics Services [NDAS]	3GPP	2022
[11] TS 28.530	Management and orchestration; Concepts, use cases and requirements	3GPP	2018
[12] TS 28.533	MANO framework. Includes Management Data Analytics Service (MDAS)	3GPP	2022
[13] TS 28.545	Fault Supervision (FS)	3GPP	2021
[14] TS 28.552	5G; Management and orchestration; 5G performance measurements	3GPP	2022
[15] TS 28.554	Management and orchestration; 5G end to end Key Performance Indicators (KPI)	3GPP	2022
[16] AI Act	Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Act	EC	2021

[17] RFC 8911	Performance Metrics Registry	IANA	2023
[18] RFC 2498	IPPM Metrics for Measuring Connectivity	IETF	1999
[19] RFC 2544	Benchmarking Methodology for Network Interconnect Devices	IETF	1999
[20] RFC 2681	A Round-trip Delay Metric for IPPM	IETF	1999
[21] RFC 3261	SIP: Session Initiation Protocol	IETF	2002
[22] RFC 3376	Internet Group Management Protocol, Version 3	IETF	2002
[23] RFC 3393	IP Packet Delay Variation Metric for IP Performance Metrics (IPPM)	IETF	2002
[24] RFC 3611	RTP Control Protocol Extended Reports (RTCP XR)	IETF	2003
[25] RFC 4656	A One-way Active Measurement Protocol (OWAMP)	IETF	2006
[26] RFC 4737	Packet Reordering Metrics	IETF	2006
[27] RFC 5357	A Two-Way Active Measurement Protocol (TWAMP)	IETF	2008
[28] RFC 6534	Loss Episode Metrics for IP Performance Metrics (IPPM)	IETF	2012
[29] RFC 6673	Round-Trip Packet Loss Metrics	IETF	2016
[30] RFC 7398	A Reference Path and Measurement Points for Large-Scale Measurement of Broadband Performance	IETF	2014
[31] RFC 7497	Rate Measurement Test Protocol Problem Statement and Requirements	IETF	2015
[32] RFC 7679	A One-Way Delay Metric for IP Performance Metrics (IPPM)	IETF	2016
[33] RFC 7680	A One-Way Loss Metric for IP Performance Metrics (IPPM)	IETF	2016
[34] RFC 7799	Active and Passive Metrics and Methods (and everything in-between, or Hybrid),	IETF	2016
[35] RFC 8762	Simple Two-Way Active Measurement Protocol (STAMP)	IETF	2021
[36] RFC 8914	Extended DNS Errors	IETF	2020
[37] RFC 9544	Precision Availability Metrics (PAMs) for Services Governed by Service Level Objectives (SLOs)	IETF	2024

[38] ATIS-0800008.v002	Quality of Service Metrics for Linear (Broadcast) IPTV	ATIS	2011
[39] ATSC A/78A	ATSC A/78A, "Transport Stream Verification," May 2007. ATIS-0800040, "	ATSC	2007
[40] ATIS-0800040	IPTV MPEG Transport Stream Monitoring	ATIS	2010
[41] ATIS-0800041	Implementers Guide to IPTV QoS Metrics	ATIS	2009
[42] GS-NFV-TST 008	Network Functions Virtualization (NFV) Release 3; Testing; NFVI Compute and Network Metrics Specification	ETSI	2020
[43] 5G-PPP	5G-PPP White Paper: Beyond 5G/6G KPIs and Target Values	5G PPP	2022
[44] 802.3.2	IEEE Standard for Ethernet - YANG Data Model Definitions	IEEE	2019
[45] BS ISO 5725	Accuracy (trueness and precision) of measurement methods and results	ISO	2023
[46] BT.1907	Objective perceptual video quality measurement techniques for broadcasting applications using HDTV in the presence of a full reference signal.	ITU-R	2012
[47] E.803	Quality of service parameters for supporting service aspects	ITU-T	2022
[48] E.804	QoS aspects for popular services in mobile networks	ITU-T	2014
[49] G.107	The E-model: a computational model for use in transmission planning	ITU-T	2015
[50] G.827	Availability performance parameters and objectives for end-to-end international constant bit-rate digital paths	ITU-T	2003
[51] G.1020	Performance parameter definitions for quality of speech and other voiceband applications utilizing IP networks	ITU-T	2006
[52] G.1028	End-to-end quality of service for voice over 4G mobile networks	ITU-T	2019
[53] G.1028.1	End-to-end quality of service for video telephony over 4G mobile networks	ITU-T	2019
[54] G.1030	Estimating end-to-end performance in IP networks for data applications	ITU-T	2014
[55] G.1031	QoE factors in web-browsing	ITU-T	2014
[56] G.1032	Influence factors on gaming quality of experience	ITU-T	2017
[57] G.1033	Quality of service and quality of experience aspects of digital financial services	ITU-T	2019
[58] G.1035	Influencing factors on quality of experience for virtual reality services	ITU-T	2021
[59] G.1036	Quality of experience influencing factors for augmented reality services	ITU-T	2022
[60] G.1050	Network model for evaluating multimedia transmission performance over Internet Protocol	ITU-T	2016

[61] G.1080	Quality of experience requirements for IPTV services	ITU-T	2008
[62] G.1091	Quality of Experience requirements for telepresence services	ITU-T	2014
[63] G.8052.1	Operation, administration, maintenance (OAM) management information and data models for the Ethernet-transport network element	ITU-T	2023
[64] J.246	Perceptual visual quality measurement techniques for multimedia services over digital cable television networks in the presence of a reduced bandwidth reference	ITU-T	2008
[65] J.247	Objective perceptual multimedia video quality measurement in the presence of a full reference	ITU-T	2008
[66] P.1201	Parametric non-intrusive assessment of audiovisual media streaming quality	ITU-T	2012
[67] P.910	Subjective video quality assessment methods for multimedia applications	ITU-T	2022
[68] G.8013/Y.1731	Operation, administration and maintenance (OAM) functions and mechanisms for Ethernet-based networks	ITU-T	2015
[69] Y.1540	Internet protocol data communication service – IP packet transfer and availability performance parameters	ITU-T	2002
[70] Y.1563	Ethernet frame transfer and availability performance.	ITU-T	2009
[71] Y.2614	Network reliability in public telecommunication data networks	ITU-T	2008
[72] MEF 38	Service OAM Fault Monitoring YANG Module	MEF	2012
[73] MEF 39	Service OAM Performance Monitoring YANG Module	MEF	2012
[74] SP 800-61	Computer Security Incident Handling Guide	NIST	2012
[75] AWS CloudWatch	AWS CloudWatch metrics	AWS	2023

3.2 Definitions

The following terminology is used throughout this Technical Report.

Term	Definition
Advantage factor	This factor represents the customer's acceptance in some decrease in quality for access advantage, e.g., mobility or connection into hard-to-reach regions.
Availability	Availability is the uptime divided by the total time.
Haptics	Technology that stimulates the senses of touch and motion.
Reliability	Reliability is the length of time the system operates without disruption.
Retainability	Retainability is the number of losses of a QoS flow per unit time.
Survival time	Survival time is the time that the application may continue without receiving an anticipated message.

3.3 Abbreviations

This Technical Report uses the following abbreviations:

Term	Definition
AGF	Access Gateway Function
AI	Artificial Intelligence
AMF	Access & Mobility Management Function
AWS	Amazon Web Services
A/V	Audio and Video
BNG	Broadband Network Gateway
CDN	Content Delivery Network
CIR	Committed Information Rate
CPE	Customer Premises Equipment.
DNS	Domain Name System
DRM	Digital Rights Management
EE	Energy Efficiency
EIR	Excess Information Rate
eMBB	Enhanced Mobile Broadband
EPON	Ethernet Passive Optical Network
FEC	Forward Error Correction
gNB	Next Generation Node B
HSI	High Speed Internet
IGMP	Internet Group Management Protocol
InP	Infrastructure Provider
IPPM	IP Performance Measurement
MOS	Mean Opinion Score
MMLU	Measuring Massive Multitask Language Understanding
MMTC	Massive Machine Type Communication
MPEG TS	Moving Pictures Expert Group Transport Stream
MTBF	Mean time between failure
MTTR	Mean time to repair
NFV	Network Functions Virtualization
NWDAF	Network Data Analytics Function
OWAMP	One-way Active Measurement Protocol
PSNR	Peak Signal to Noise Ratio
QED	Quality of Experience Delivered
QoS	Quality of Service
RTP	Real-time Transport Protocol

RTCP XR	RTP Control Protocol Extended Reports
SDN	Software Defined Network
SES	Severely Errored Seconds
STAMP	Simple Two-Way Active Measurement Protocol
TR	Technical Report
TWAMP	Two-Way Active Measurement Protocol
UDM	Unified Data Management
UDPST	UDP-based speed tests
URLLC	Ultra Reliable Low Latency Communications
VNO	Virtual Network Operator
WLAN	Wireless Local Area Network
WT	Working Text

4. Technical Report Impact

4.1 Energy Efficiency

TR-499 includes metrics related to energy efficiency, which can be used to monitor and improve energy efficiency.

4.2 Security

TR-499 includes metrics related to security violations, which can be used to monitor and improve security.

4.3 Privacy

TR-499 has no impact on privacy.

5. Introduction

Operators are now moving above simple broadband access toward deploying platforms for services delivery. Service platforms can encompass edge or cloud computing resources for service origination, and aspects such as real-time sessions, in addition to network delivery. Future services have widely different characteristics and requirements, and so they have different emphasis on service assurance and need a variety of metrics to ensure delivery.

Various sets of metrics should be used to ensure services delivery, depending on the desired operations, the entity involved, technology and service type. Use cases and example services drive the establishment of service metric sets. Also, metric sets are accompanied by a number of related dimensions, such as the statistics that are applied to each metric, the applicable network segment, etc. This is depicted in Figure 1.

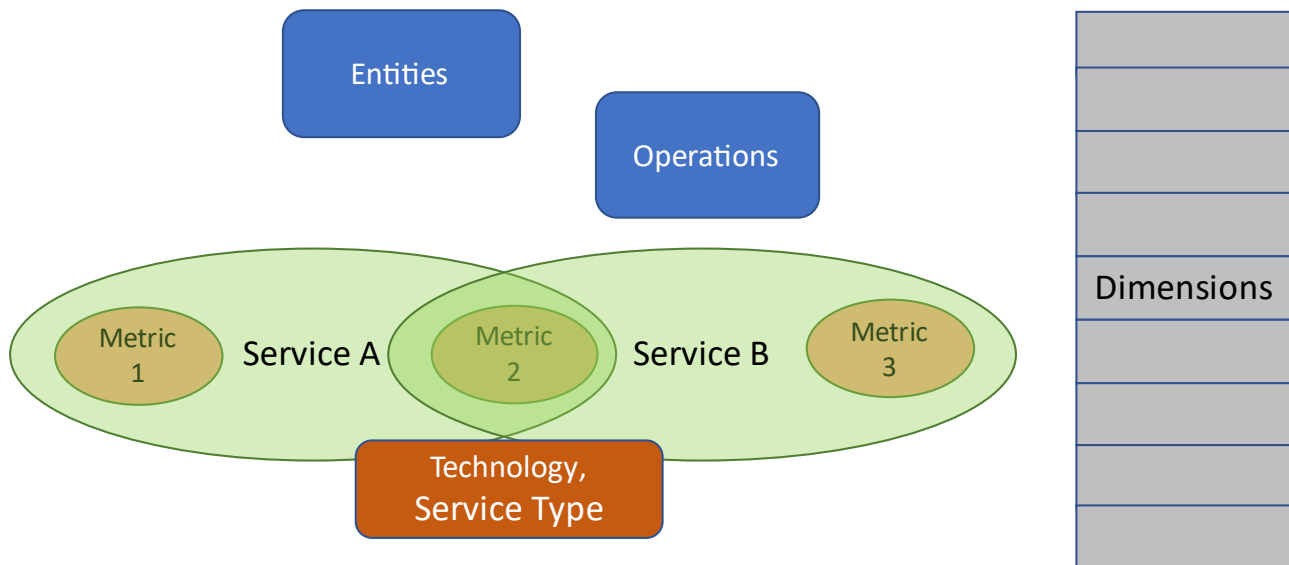


Figure 1. Metrics, services, use cases, and dimensions.

Section 6 defines the overall sets of service metrics and their definitions. Service metrics are categorized into quality layers: Transaction quality, Content quality, Flow quality, and Network quality. Section 7 defines dimensions related to service metrics. Section 8 defines sets of metrics for each example service, these metrics were collated to define the overall metrics of Section 6. Section 9 identifies use cases where service metrics can be applied. 0 summarizes many existing documents that define metrics which were used to identify the service metrics here.

6. Overall Service Metrics

The metrics of all the example services in Section 8 were collated, and common metrics across the services are identified, with references to their definitions, as shown in Table 1. Metrics with names in *purple italics* are specific to some services, while other metrics are common across all services. Tables in Section 8 also include capacity and examples of metrics.

Quality layers identify various perspectives of service delivery in which quality can be measured by means of appropriate metrics. Quality layers are:

- ♦ *Transaction Quality*: Sessions, control and management planes, and interactions between the user and the service.

- ◆ *Content Quality*: Actual content: video, audio, haptics, etc.
- ◆ *Flow Quality*: Quality per packet flow, slice or transport stream.
- ◆ *Network Quality*: Performance of the underlying network.

Different quality layers are exposed at different points in the service delivery chain. Content quality is generally only accessible at the application endpoints, service origination and terminal. Transaction quality is generally accessible only at the terminal, service origination and other control-plane functions. Flow and network quality may be available throughout the network, it is also possible for a special-purpose probe to access any quality layer at any point in the network.

Table 1. Overall Service Metrics

Metric	Definition
TRANSACTION QUALITY	
Session set-up time	The time from a request to initiate a session until it is fully established and running. 3GPP TS 28.554 [15] has specific occurrences, e.g., PDU session establishment time of a network slice.
Transaction delays	The time from when a transaction is initiated until its completion. ATIS-0800008 [38] defines Channel Change Delay, including IGMP leave and join times. Also, ITU-T G.1080 [61].
Number of successful service requests	The number of times a service is successfully fulfilled. 3GPP TS 28.554 [15] has specific occurrences, e.g., Registration success rate of one single network slice.
Number of sessions or subscribers	Count of the number of established sessions or connected customers. 3GPP TS 28.554 [15] has specific occurrences, e.g., Registered subscribers of network through UDM.
Interactions between the user and the service	Widely varying, includes service portal interactions, etc. ITU-T E.803 [47] has many occurrences such as complaint management, and ATIS-0800008 [38] has occurrences including Electronic Program Guide (EPG) functionality.
Control and management planes	Wide range, salient aspects depend on the service. Can include DNS errors [36], DRM Errors [41], SIP stats [21], Management VLAN connectivity, DHCP, and server selection.
<i>Energy efficiency and consumption</i>	3GPP TS 28.554 [15] defines multiple KPIs for energy efficiency and energy consumption. Also, the durations and performance of power saving modes (PSM): data rate in PSM, time to exit PSM, percent time in PSM, lost packets during PSM, failures to enter/exit PSM .
<i>Fault events</i>	Number of fault events, which can be stratified to different types of faults. TR-181 Device.FaultMgmt. [3]
<i>Security violations</i>	Numbers of security incidents, threats, and breaches. Computer Security Incident Handling Guide (nist.gov) [74], Appendix B Incident-Related Data Elements. ITU-T XSTR-SEC-MANUAL.
CONTENT QUALITY	
Display and output quality (video, audio, A/V synch, haptics,...)	Video picture quality and Mean Opinion Score (MOS) estimates: TR-126 [1], ITU-T J.247 [65], ITU-T J.246 [64], ITU-T P.1201 [66], ITU-T BT.1907; video peak signal to noise ratio (PSNR) [46], [38], VQEG https://www.vqeg.org/publications-and-software/publications/ , frame rate, resolution. Audio quality: ITU-T G.107 [49], G.1020 [51]. Other services quality: ITU-T G.1028-1035, G.1050 [52]-[60].

Drop-outs, lost samples	The number of lost data samples, or occurrences when the service was interrupted. TR-126 [1], ITU-T P.910 [67], ITU-T G.803 [47]
Availability and reliability	Availability is the uptime divided by the total time. Reliability is the length of time the system operates without disruption. ITU-T G.827 [50], ITU-T E.804 [48] RFC 9544 [37], Precision Availability Metrics for Services Governed by Service Level Objectives (SLOs).
Artifacts	Numbers of perceptible or imperceptible abnormalities in picture, audio or other presentation. TR-126 [1], ITU-T P.910 [67], ITU-T J.247 [65], ITU-T J.246 [64].
Survival time	Survival time is the time that the application may continue without receiving an anticipated message. 5G PPP [43].
Content origination: compute/cloud metrics	NFV infrastructure metrics ETSI GS NFV-TST 008 [42] , AWS CloudWatch metrics [75]. Uptime, availability, MTBF, MTTR, CPU, Memory, Disk, Requests per minute Time to acknowledge or respond, Latency, Error rate, Swap usage, Throughput (transactions per second), Scalability, Load balancing.
Terminal (headset) status	The status and resource usage within a service terminal device. TR-181 [3], ProcessStatus , MemoryStatus , TemperatureStatus .
<i>Real-time interactivity and delay</i>	Performance and delay of service interactions while a service is in use. ITU-T G.1031 [55] Web, G.1032 [56] Gaming, G.1035 [58] VR, G.1036 [59] AR, G.1091 [62] Telepresence.
<i>Parameter resolution and accuracy</i>	Accuracy is the difference between a reported parameter value and its actual value. Resolution is the smallest increment between quantized samples in the represented reporting format. BS ISO 5725 [45], Mean squared error, absolute error.
<i>Interactivity, Useability</i>	For AI systems, ability to simulate human behavior, input and output natural language.
<i>Model accuracy</i>	For AI models. Error in output; including Mean squared error (MSE), absolute error, accuracy, robustness, regression R² , false positives, false negatives, precision, recall, confusion matrix (see Note). Utility function. Benchmarking. ITU-T Artificial Intelligence Quotient (E.AIQ), MMLU (for language models).
<i>Traceability, Regulatory</i>	For AI systems. includes transparency, documentation, conformity assessment [16]. Model bias/fairness [16]. Performance of guardrails limiting criminal or unethical use, explainability, explainable AI (XAI).
FLOW QUALITY	
Data rate	Measured throughput TR-471 [9], TR-143 [2]; Allocated throughput CIR and EIR [72][73]; IETF RFC 7497 [31]. RFC 2544 [19], RFC 5357 [27] RFC 8762 [35].
Delay	Delay or latency, QED TR-452 .1 [6], TR-452.2 [7], TR-452.5 [8]. RFC 7679 [32], RFC 2681 [20], IANA Performance Metrics [17].
Loss rates	Packet loss rate (PLR) after all error correction, counts of retransmissions, error rate before error correction, code violations. RFC 7680 [33], RFC 6673 [29], RFC 6534 [28], RFC 2544 [19], IANA Performance Metrics [17].
Jitter	The variation in the arrival times of successive packets. IP Packet Delay Variation (IPDV), RFC 3393 [23].
Numbers of flows or slices	Numbers of flows, streams, or slices. 3GPP TS 28.554 [15] has specific occurrences, e.g., number of PDU sessions of network slice.
Retainability	Retainability is how often an end-user abnormally loses a QoS flow. ITU-T E.804 [48]. 3GPP TS 28.554 [15], Retainability KPI. ITU-T Y.1540 [69], RFC 9544 [37].
<i>Impact from other flows</i>	Non-ideal slice isolation results in be adverse impact by resource usage of other slices or network utilization. Similarly, flows configured or allocated QoS levels

	through the network environment may be adversely impacted by resource usage of other flows or network utilization. 3GPP TS 28.530 [11].
<i>MPEG Transport Stream (TS) or RTCP XR stats, IGMP join/leave times</i>	ATSC A/78A [39] and ATIS-0800040 [40] define MPEG TS stats. RFC 3611 [24] defines RTCP XR stats, and was extended in RFC 7003, RFC 7005, RFC 7243, RFC 7244, and RFC 7294. IGMP is defined in RFC 3376 [22]. ATIS-0800008 [38] and ITU-T G.1080 [61] describe IGMP join and leave times.
<i>Routes</i>	Route flapping (changing routes) and its impact, throughput variations, out-of-order packets, RFC 4737 [26].
NETWORK QUALITY	
Data rate	Measured throughput TR-471 [9], TR-143 [2]; Allocated throughput CIR and EIR [72][73]; IETF RFC 7497 [31]. RFC 2544 [19], RFC 5357 [27] RFC 8762 [35] RFC 5357 [27].
Delay	Delay or latency, QED TR-452 .1 [5], TR-452.2 [7], TR-452.5 [8]. RFC 7679 [32],. RFC 2681 [20].
Loss rates	Packet loss rate (PLR) after all error correction, counts of retransmissions, error rate before error correction, code violations. RFC 7680 [33], RFC 6673 [29], RFC 6534 [28], RFC 2544 [19].
Jitter	The variation in the arrival times of successive packets. IP Packet Delay Variation (IPDV), RFC 3393 [23].
Availability and reliability	Availability is the uptime divided by the total time. Reliability is the length of time the system operates without disruption. ITU-T G.827 [50], ITU-T E.804 [48], ITU-T Y.1540 [69], ITU-T Y.1563 [70], ITU-T Y.2614 [71].
Performance counters	Counters such as BytesSent, BytesReceived, PacketsSent, PacketsReceived, ErrorsSent, ErrorsReceived, DiscardPacketsSent, DiscardPacketsReceived, RetransCount. TR-181 [3] Device.IP.Interface.{i}.Stats .
Network equipment status	Status, resource usage, temperature, etc., of network elements. TR-181 [3], ProcessStatus , MemoryStatus , TemperatureStatus .
<i>Connections</i>	Total numbers of connections and dropped connections. RFC 2498 [18]. [71] [68]. draft-ietf-ippm-connectivity-monitoring-06, A Connectivity Monitoring Metric for IPPM, 2023. Ethernet OAM; ITU-T G.8052.1 [63], MEF 38 and 39 [72][73].

Note:

Percent accuracy = 100*number of correct predictions divided by the total number of predictions.

Precision = true positives / (true positives + false positives).

Recall = true positives / (true positives + false negatives).

The confusion matrix shows estimated positives and negatives in columns, and true positives and negatives in rows, with correct estimations on the diagonal.

7.Overall Service Metric Dimensions

A service metric, or a set of service metrics, is accompanied by a number of related dimensions, such as the statistics used to represent each metric, the measurement point, the measurement time, etc. The metric dimensions of Section 8 were collated, and common metric dimensions across all services are identified as shown in Table 2.

Table 2. Overall Service Metric Dimensions

Statistic: minimum, maximum, average, percent points, histogram, time series, counters
Time: timestamps, measurement interval (duration), tick interval (between recording samples) Reporting: periodic, on demand, threshold-based
Measurement point: E2E, network segment, network node, terminal, content origination, test device, probe
Direction, downstream, upstream, two-way
Total available, per QoS class, per flow, per sub-flow, per service instance
Data aggregation and presentation: overall, per network segment, per user, per device, geographic area, domain
Domain: content provider, service provider, network provider, consumer; Infrastructure Provider (InP), Virtual Network Operator (VNO)
Resolution and accuracy
Measured usage, or capacity, or required or advertised. Percent utilization.
Measurement layer: transaction, content, flow, network; OSI layer
Advantage factor: This factor represents the customer's acceptance in some decrease in quality for access advantage, e.g., mobility or connection into hard-to-reach regions, display terminal size
Shared resource vs. per instance: e.g., DNS load balancing, slice isolation.

8. Example Services and their Metrics

Exemplary services are considered here, and applicable service metrics and dimensions are identified for each service.

8.1 Service Metrics example: XR Rendering

This service is rendering XR content across the network. As the user moves the XR headset, motion information is sent up to the edge compute which renders the video and sends it back down to the headset. The apparent delay of this loop must be extremely low (~ 2 millisecc), although somewhat higher delays can be concealed by timewarp (a term used in XR rendering for motion compensation) or motion prediction.

Needs for delivering this service include:

- Extremely low latency from edge compute to headset (~ 2 millisecc), both upstream and downstream
- 30-200 Mbps throughput downstream

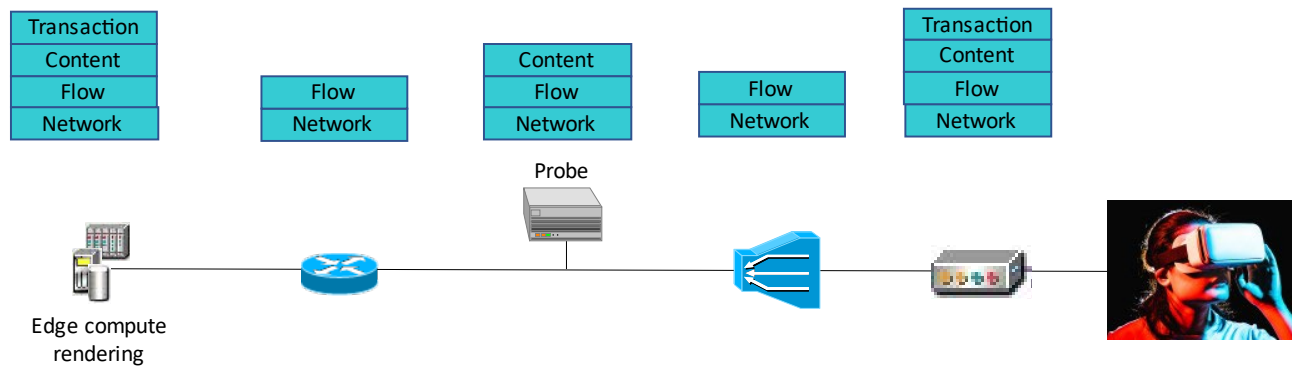


Figure 2 Quality layers for XR Rendering.

The figure above depicts quality layers for metrics to ensure service delivery and where they may be measured. Sets of service metrics for this service are shown in the table below for each quality layer.

Quality layers identify various perspectives of service delivery in which quality can be measured by means of appropriate metrics. Quality layers are:

- ◆ *Transaction Quality*: Sessions, control and management planes, and interactions between the user and the service.
- ◆ *Content Quality*: Actual content: video, audio, haptics, etc.
- ◆ *Flow Quality*: Quality per packet flow, slice or transport stream.
- ◆ *Network Quality*: Performance of the underlying network.

Metric	Capacity	Example(s)
Transaction Quality		
Session set-up time		Time from request to service start
Transaction delays		Time to switch between "worlds"
Number of successful service requests	Maximum supported number of service requests	
Number of sessions or subscribers	Maximum supported number of sessions or subscribers	
Interactions between the user and the service		Service portal interactions, service sign-on, service modifications
Control and management planes		Edge server selection, Management VLAN

Content Quality		
Display and output (video, audio, A/V synch, haptics,...) quality	Quality before network transmission	Service Mean Opinion Score (MOS), picture quality, audio quality, haptics quality, A/V/haptics synch (< 10 millisecc), rendering, user dizziness
Drop-outs		
Availability and reliability		Availability to host sessions and reliability to continue sessions.
Artifacts		Impairment scales
Survival time		The time that the application may continue without receiving an anticipated message
Content origination: compute/cloud metrics	Capacity	Time to render, CPU, memory, disk, compute time, number of supported sessions [75]
Terminal (headset) status	Headset capacity	Equipment status (up/down, etc.), CPU, memory, network, buffer fill, energy consumption, firmware version, decoder
Real-time interactivity and delay		VR “motion-to-photon” latency and QED, (P90 < 2 millisecc; P99.9 < 5 millisecc; P100 < 10 millisecc), Timewarp performance
Flow Quality		
Data rate	Network capacity, flow assignment	CIR/EIR, 30-200 Mbps throughput downstream
Delay	Buffer sizes, application tolerance, or quality loss	QED
Loss rates	Application tolerance	Packet Loss Rate (PLR)
Jitter		IP Packet Delay Variation (IPDV)
Numbers of flows or slices	Maximum supported number of slices or flows	Number of slices or flows, resource utilization, requests (success/fail)
Retainability		Number of losses of a QoS flow per unit time
Impact from other flows	Slice isolation	
Network Quality		
Data rate	Network capacity	CIR/EIR, throughput
Delay	Buffer sizes, application tolerance, or quality loss	QED
Loss rates		Packet Loss Rate (PLR)

Jitter		IP Packet Delay Variation (IPDV)
Availability and reliability		Availability, reliability, Severely Errored Seconds (SES), Mean Time Between Failures (MTBF), Mean Time to Repair (MTTR)
Performance counters		BytesSent, BytesReceived, PacketsSent, PacketsReceived, ErrorsSent, ErrorsReceived, DiscardPacketsSent, DiscardPacketsReceived, RetransCount
Network equipment status		Status, temperature, CPU and memory utilization

8.1.1 Applicable Service Metric Dimensions

- Statistic: minimum, maximum, average, percent points, histogram, time series
- Time: timestamps, measurement interval (duration), tick interval (between recordings)
- Measurement point
- Direction, downstream, upstream, two-way,
- Total available or per flow

8.2 Service Metrics example: Multiple IoT Sensors

Multiple IoT sensors may also be referred to as massive Machine-Type Communications (mMTC). This service excludes high-bandwidth sensors such as security cameras.

Needs for delivering this service include:

- Support for many connections
- Optional proxy function
- Energy savings

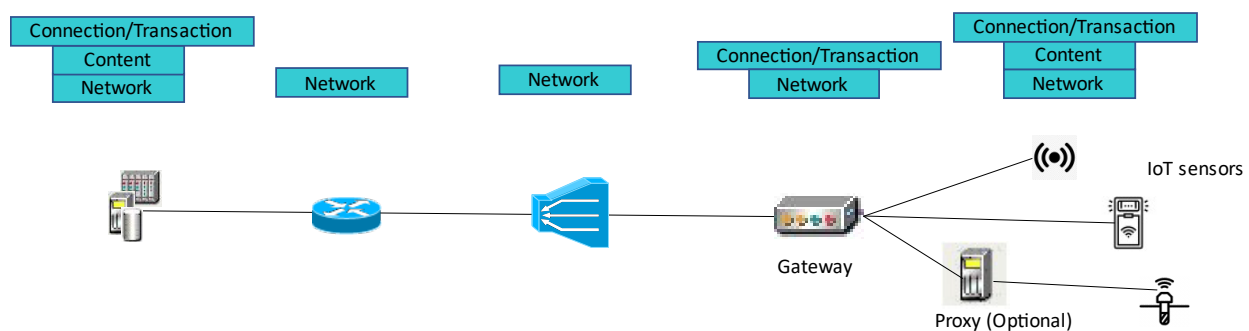


Figure 3 Quality layers for Multiple IoT sensors.

The figure above depicts quality layers for metrics to ensure service delivery and where they may be measured. Sets of service metrics for this service are shown in the table below for each quality layer.

Metric	Capacity	Example(s)
Connection and Transaction Quality		
Session set-up time		Time from request to connection of device
Numbers of sessions		Numbers of sessions current/successful/failed
Interactions with the sensor service		Service sign-on, service modifications, on-demand data requests
Energy efficiency and energy consumption		Energy efficiency (bits/Joule), Energy consumption (Watt hours). Also, performance of power saving modes (PSM) data rate in PSM, time to exit PSM, percent time in PSM, lost packets during PSM, failures to enter/exit PSM
Control and management planes		Data aggregation, storage, and presentation. Cloud metrics, memory usage
Fault Events		Missed security alarm
Content Quality		
Lost samples		Total number of expected samples - Total # received samples, loss rate
Availability and reliability		Availability of sensor or domain to provide measurements. Reliability to continue reading.
Survival time		The time that the application may continue without receiving an anticipated message
Content origination: compute/cloud metrics		Cloud Metrics [75]
Terminal (sensor) status		Equipment status (up/down, etc.), energy consumption, firmware version
Delay between samples		QED
Parameter resolution and accuracy		Mean squared error, absolute error, statistics
Network Quality		
Delay		QED
Loss rates		Packet Loss Rate (PLR)

Availability and reliability		Network availability, reliability, Severely Errored Seconds (SES), Mean Time Between Failures (MTBF), Mean Time to Repair (MTTR)
Performance counters		BytesSent, BytesReceived, PacketsSent, PacketsReceived, ErrorsSent, ErrorsReceived, DiscardPacketsSent, DiscardPacketsReceived, RetransCount
Network equipment status		Status, temperature, CPU and memory utilization
Connections	Max supportable number of connections	Current number of connections

8.2.1 Applicable Service Metric Dimensions

- Statistic: minimum, maximum, average, percent points, histogram, time series
- Time: timestamps, measurement interval (duration), tick interval (between recordings)
 - Collection times: periodic, notification (including threshold based), on-demand
- Measurement point
 - Direct to sensor, or to sensor via proxy
- Data aggregation and presentation: per sensor, per domain, per premises

8.3 Service Metrics example: High-Speed Internet (HSI) access

Needs for delivering this service include:

- High speed
- Low delay
- High availability

Here network performance is paramount, although some higher-layer functions such as DNS are also important. Sets of service metrics for this service are shown in the table below for each quality layer.

Metric	Capacity	Example(s)
Connection and Transaction Quality		
Connection times, session set-up time		Service sign-on time from modem boot
DNS availability		DNS is available and running properly

Network Quality		
Data rate	Max data rate	Available throughput, may need to measure very high speeds with UDPST
Delay		QED
Loss rates		Packet Loss Rate (PLR)
Jitter		IP Packet delay variation (IPDV)
Availability and reliability		Network availability, reliability, Severely Errored Seconds (SES), Mean Time Between Failures (MTBF), Mean Time to Repair (MTTR)
Performance counters		BytesSent, BytesReceived, PacketsSent, PacketsReceived, ErrorsSent, ErrorsReceived, DiscardPacketsSent, DiscardPacketsReceived, RetransCount
Network equipment status		Status, temperature, CPU and memory utilization

8.3.1 Applicable Service Metric Dimensions

- Statistic: minimum, maximum, average, percent points, histogram, time series
- Time: timestamps, measurement interval (duration), tick interval (between recordings)
 - Collection times: periodic, notification (including threshold based), on-demand
- Measurement point
 - Network node, network segment
 - Across NNI for Fixed-Access Network (FANS) scenarios
 - Measurement direction
- Data aggregation and presentation: per end-user or per network segment

8.4 Service Metrics example: Entertainment video

This service provides entertainment video. It includes linear IPTV delivered using multicast as well as unicast streaming. Most video is now delivered over HTTP/HTTPS using streaming techniques similar to MPEG DASH and HTTP live streaming (HLS), which caches and sends chunks of video data.

Needs for delivering this service include:

- High speed
- Bounded jitter for linear TV

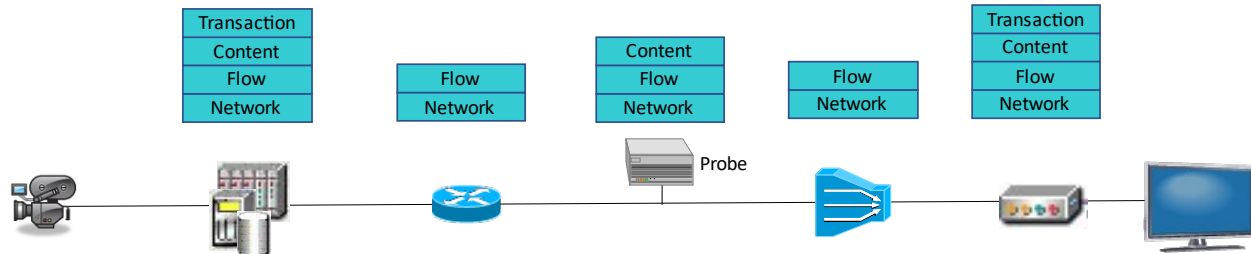


Figure 4 Quality layers for entertainment video.

The figure above depicts quality layers for metrics to ensure service delivery and where they may be measured. Sets of service metrics for this service are shown in the table below for each quality layer.

Metric	Capacity	Example(s)
Transaction Quality		
Session set-up time		Time from request to service start, boot time
Transaction delays		Time from video request to decoding streams. Time to change channels (zap time)
Number of successful service requests	Maximum supported number of service requests	
Number of sessions or subscribers	Maximum supported number of sessions or subscribers	
Interactions between the user and the service		Service portal interactions, service sign-on, service modifications, Electronic Program Guide (EPG) functionality
Control and management planes		Digital Rights Management (DRM) errors, Management VLAN
Content Quality		
Video and audio quality	Quality before network transmission	Video and audio Mean Opinion Score (MOS), frame rate, resolution, picture quality, PSNR, audio quality, A/V synch (lip synch)
Drop-outs		Video and audio dropouts
Availability and reliability		Availability to start sessions, reliability to continue sessions
Artifacts		Impairment scales. Perceptible and hidden video artifacts
Survival time		The time that decoding and display may continue without receiving an anticipated message
Content origination: compute/cloud metrics	Capacity	Time to encode or stream, CPU, memory, disk, compute time, number of supported sessions, etc. [75]

		Local content insertion
Terminal (set-top or decoder) status	Terminal device capacity	Decoder buffer overflow/underrun. Equipment status (up/down, etc.), CPU, memory, network, energy consumption, firmware version, decoder
Flow (or Stream) Quality		
Data rate	Network capacity, flow assignment	Downstream. Constant (CBR) and Variable (VBR) bit rate
Delay	Buffer sizes, application tolerance, or quality loss	QED
Loss rates	Decoder tolerance	Packet Loss Rate (PLR). Burst loss: burst frequency, burst length, loss rate during burst. Before and after error correction
Jitter		IP Packet delay variation (IPDV)
Numbers of flows or slices	Maximum supported number of video and audio flows	Number of flows, resource utilization, requests (success/fail)
Retainability		Number of losses of a QoS flow per unit time
MPEG Transport Stream (TS) or RTCP XR stats, IGMP join/leave		Lost timestamps/timing errors. Per packet identifier (PID). A/V synch, loss of mapping tables
Network Quality		
Data rate	Network capacity	
Delay	Buffer sizes, application tolerance, or quality loss	QED, IGMP leave/join time
Loss rates		Packet Loss Rate (PLR) Burst loss: burst frequency, burst length, loss rate during burst
Jitter		IP Packet delay variation (IPDV)
Availability and reliability		Availability, reliability, Severely Errored Seconds (SES), Mean Time Between Failures (MTBF), Mean Time to Repair (MTTR)
Performance counters		BytesSent, BytesReceived, PacketsSent, PacketsReceived, ErrorsSent, ErrorsReceived, DiscardPacketsSent, DiscardPacketsReceived, RetransCount
Network equipment status		Status, temperature, CPU and memory utilization

8.4.1 Applicable Service Metric Dimensions

- Statistic: minimum, maximum, average, percent points, histogram, time series

- Time: timestamps, measurement interval (duration), tick interval (between recordings)
- Measurement point
- Total available, or per flow
 - Per sub-flow: video (L/R video if 3D), audio (L/R), secondary audio, per packet identifier (PID)
- Resolution and accuracy
- Domain: content provider, service provider, network provider, consumer

8.5 Video Teleconferencing

This use case is interactive video teleconferencing among multiple participants. The video, audio and screen share that are uploaded by participants are amalgamated by a cloud server and sent back down. Both upstream and downstream are used.

Needs for delivering this service include:

- Bounded delay
- High availability and reliability

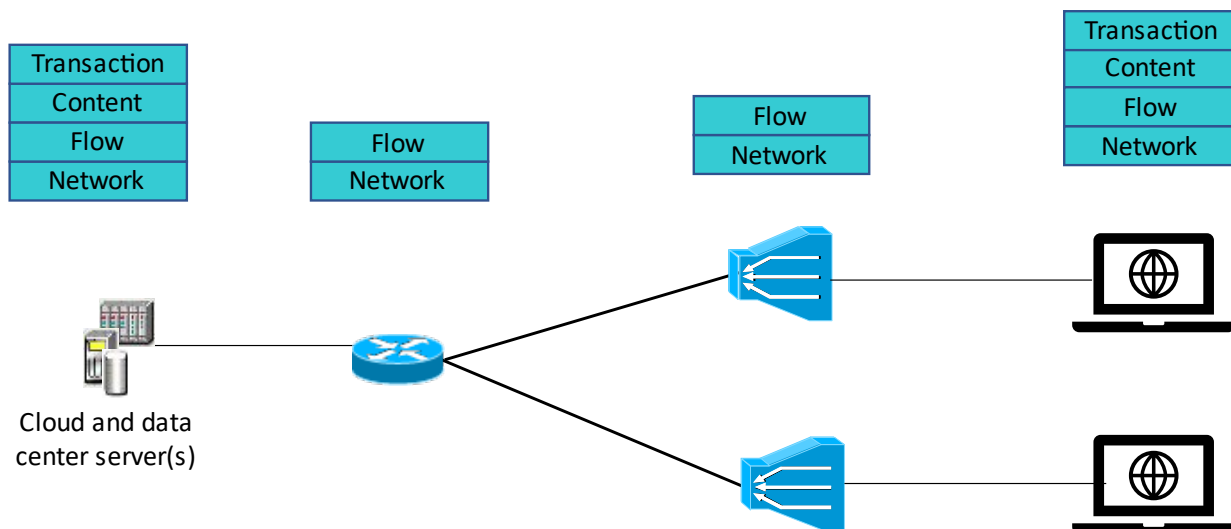


Figure 5 Quality layers for video teleconferencing.

The figure above depicts quality layers for metrics to ensure service delivery and where they may be measured. Sets of service metrics for this service are shown in the table below for each quality layer.

Metric	Capacity	Example(s)
Transaction Quality		
Session set-up time		Time from request to service start
Transaction delays		Time to put teleconference controls into effect

Number of successful service requests	Maximum supported number of service requests	
Number of sessions or subscribers	Maximum supported number of sessions or subscribers	
Interactions between the user and the service		Service portal interactions, service sign-on, setting up meetings, sharing content
Control and management planes		Connection process, Management VLAN connectivity, DHCP, server selection
Security violations		Numbers of security incidents, threats, and breaches.
Content Quality		
Display and audio quality	Quality before network transmission	Video and audio Mean Opinion Score (MOS), frame rate, resolution, picture quality, PSNR, audio quality, A/V synch (lip synch)
Drop-outs		Video and audio dropouts
Availability and reliability		Availability to connect to server(s) and start sessions. Session continuation reliability
Artifacts		Impairment scales. Perceptible and hidden video artifacts
Survival time		The time that decoding and display may continue without receiving an anticipated message
Content origination: compute/cloud metrics	Capacity	Cloud metrics, CPU, memory, disk, compute time, number of simultaneous users [75]
Terminal status	Terminal device capacity	Decoder buffer fill, buffer overflow/underrun. Equipment status (up/down, etc.), CPU, memory, network, energy consumption, firmware version
Real-time interactivity		Delay from a user's input to another user's display
Flow (or Stream) Quality		
Data rate	Network capacity, flow assignment	Downstream and upstream. HTTP tunnels
Delay		QED. Time to put teleconference controls into effect
Loss rates	Decoder tolerance	Packet Loss Rate (PLR). Burst loss: burst frequency, burst length, loss rate during burst
Jitter		IP Packet delay variation (IPDV)
Numbers of flows or slices	Maximum supported number of video and audio flows	Number of flows, resource utilization

Retainability		Number of losses of a QoS flow per unit time
Routes		Route flapping, throughput variations, out-of-order packets
Network Quality		
Data rate	Network capacity	
Delay	Buffer sizes, application tolerance, or quality loss	QED
Loss rates		Packet Loss Rate (PLR) Burst loss: burst frequency, burst length, loss rate during burst
Jitter		IP Packet delay variation (IPDV)
Performance counters		BytesSent, BytesReceived, PacketsSent, PacketsReceived, ErrorsSent, ErrorsReceived, DiscardPacketsSent, DiscardPacketsReceived, RetransCount
Availability and reliability		Availability, reliability, Severely Errored Seconds (SES), Mean Time Between Failures (MTBF), Mean Time to Repair (MTTR)
Network equipment status		Status, temperature, CPU and memory utilization

8.5.1 Applicable Service Metric Dimensions

- Statistic: minimum, maximum, average, percent points, histogram, time series
- Time: timestamps, measurement interval (duration), tick interval (between recordings)
- Measurement point
- Direction, downstream, upstream, two-way,
- Total available or per flow

8.6 Ultra Reliable Low Latency Communications (URLLC)

Ultra Reliable Low Latency Communications (URLLC) is a class of services, including machine type communication (MTC), autonomous vehicles, automated traffic control, Industry 4.0, deterministic networking, robotics, teleoperation, and tactile internet.

Needs for delivering this service include:

- Ultra-low latency
- High availability and reliability

Sets of service metrics for this service are shown in the table below for each quality layer.

Metric	Capacity	Example(s)
Transaction Quality		
Session set-up time		Time from request to service start
Transaction delays		Interaction times, time from control input until feedback displayed to user.
Number of successful service requests	Total number of service requests	
Number of sessions or subscribers	Maximum supported number of sessions or subscribers	
Interactions between the user and the service		Service sign-on, service modifications
Security violations		Numbers of security incidents, threats, and breaches.
Content (or user feedback) Quality		
Drop-outs		Video and audio and haptics dropouts
Availability and reliability	~ 99.999% availability	Availability to start sessions, reliability to continue sessions
Controlled equipment		Controlled system performance and equipment status
Flow Quality		
Data rate	Network capacity, flow assignment	Downstream and upstream
Delay	~ 1 millisecond	QED
Loss rates		Packet Loss Rate (PLR).
Jitter		IP Packet delay variation (IPDV)
Numbers of flows or slices	Maximum supported number of video and audio flows	Number of flows, resource utilization, requests (success/fail)
Retainability		Number of losses of a QoS flow per unit time
Network Quality		
Data rate	Network capacity	Downstream and upstream
Delay	~ 1 millisecond	QED
Loss rates		Packet Loss Rate (PLR)
Jitter		IP Packet delay variation (IPDV)

Availability and reliability	~ 99.999% availability	Availability, reliability, Severely Errored Seconds (SES), Mean Time Between Failures (MTBF), Mean Time to Repair (MTTR)
Performance counters		BytesSent, BytesReceived, PacketsSent, PacketsReceived, ErrorsSent, ErrorsReceived, DiscardPacketsSent, DiscardPacketsReceived, RetransCount
Network equipment status		Status, availability, temperature, CPU and memory utilization

8.6.1 Applicable Service Metric Dimensions

- Statistic: minimum, maximum, average, percent points, histogram, time series
- Time: timestamps, measurement interval (duration), tick interval (between recordings)
- Measurement point
- Resolution and accuracy

8.7 Network support for Speed Tests

This is a somewhat specialized use case but is important since consumers often use speed tests to gauge services. It simply runs from a user device to a speed test server.

Network Quality		
Metric	Capacity	Example(s)
Data rate	Network capacity	Tests for high data rates may require UDPST
Delay		QED

8.7.1 Applicable Service Metric Dimensions

- Statistic: minimum, maximum, average, percent points, histogram, time series

8.8 Service Metrics example: AI as a Service (AlaaS)

Artificial Intelligence as a Service (AlaaS) is a cloud-based offering that provides access to AI software, tools, and capabilities on a subscription basis. Hyperscalers and others are now offering AlaaS via APIs, digital assistants/bots, pre-built models, and platforms that support AI.

Needs for delivering this service include:

- Low latency

- Accurate AI models and algorithms
- Rapid training
- Human-like responses (for some applications)

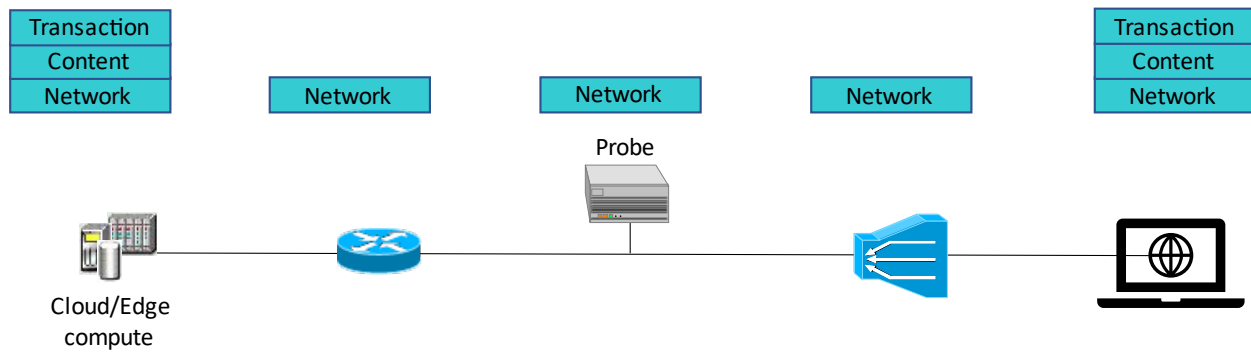


Figure 6 Quality layers for AlaaS.

The figure above depicts quality layers for metrics to ensure service delivery and where they may be measured. Sets of service metrics for this service are shown in the table below for each quality layer.

Metric	Capacity	Example(s)
Transaction Quality		
Transaction delays		Time for model to output Model training time
Number of successful service requests	Total number of service requests	
Number of sessions or subscribers	Maximum supported number of sessions or subscribers	
Interactions between the user and the service		Service portal interactions, service sign-on, service modifications.
Control and management planes		Server selection, cloud load balancing
Content Quality		
Interactivity, Useability		Ability to simulate human behavior, input and output natural language, Turing test.
Model accuracy		Error in output; Mean squared error (MSE), absolute error, accuracy, robustness, regression R² , false positives, false negatives, precision, recall confusion matrix (see Note). Hallucinations. Utility

		function. Benchmarking. ITU-T Artificial Intelligence Quotient (E.AIQ), MMLU (for language models).
Traceability, Regulatory		Transparency, documentation, conformity assessment [16]. Model bias/fairness [16]. Performance of guardrails limiting criminal or unethical use, explainability, explainable AI (XAI).
Network Quality		
Data rate	Network capacity	CIR/EIR, throughput
Delay	Buffer sizes, application tolerance, or quality loss	QED
Loss rates		Packet Loss Rate (PLR)
Jitter		IP Packet Delay Variation (IPDV)
Availability and reliability		Availability, reliability, Severely Errored Seconds (SES), Mean Time Between Failures (MTBF), Mean Time to Repair (MTTR)
Performance counters		BytesSent, BytesReceived, PacketsSent, PacketsReceived, ErrorsSent, ErrorsReceived, DiscardPacketsSent, DiscardPacketsReceived, RetransCount
Network equipment status		Status, temperature, CPU and memory utilization

Note:

Percent accuracy = $100 \times \text{number of correct predictions} / \text{the total number of predictions}$.

Precision = $\text{true positives} / (\text{true positives} + \text{false positives})$.

Recall = $\text{true positives} / (\text{true positives} + \text{false negatives})$.

The confusion matrix shows estimated positives and negatives in columns, and true positives and negatives in rows, with correct estimations on the diagonal.

8.8.1 Applicable Service Metric Dimensions

- Statistic: minimum, maximum, average, percent points, histogram, time series
- Time: timestamps, measurement interval (duration), tick interval (between recordings)

8.9 Peer-to-Peer file sharing

Peer-to-peer applications such as BitTorrent were very popular for distributing music and movies and are now used for large file transfers.

BitTorrent works by splitting files into small pieces and distributing them among many peers, who can then exchange them with each other. The BitTorrent client contacts a "tracker" specified in the .torrent file. The tracker is a special server that keeps track of the connected computers in a swarm so they can connect to each other.

The use of metrics for peer-to-peer file sharing is questionable, so this service is not explored further here.

8.10 Security cameras

Somewhat similar to Entertainment Video, but the video content flows upstream.

8.11 Gaming

Similar to URLLC; but with somewhat relaxed requirements. Also similar to video conferencing but with somewhat tighter delay requirements.

8.12 Web Browsing

Generally, just HSI plus there is a need for (possibly many) functioning server(s)/cloud with many short-lived network connections.

Appendix I. Summary of existing standards to apply to service metrics

This section shows lists of metrics as defined or included in pertinent reference standards and documents from BBF, 3GPP, ATIS, ETSI, IETF, MEF, and others. Note that some metrics are duplicated across different standards, and there are metrics that may not apply to services delivery. In addition, there are a number of other relevant standards some of which do not add to the metrics already listed here.

I.1 BBF TR-304

TR-304 [4] defines a detailed architecture for network performance measurement components. It also lists a number of metrics for access networks, which are listed here.

Table 3. [TR-304 Table 2] – Static Access Service Attributes

Access Service Provider IANA Enterprise Number	<signed integer>	IANA Enterprise Number of the Access Service Provider. IANA numbers listed at http://www.iana.org/assignments/enterprise-numbers .
Access Service Provider name	<string>	Name of the Access Service Provider Human readable non-standardized format.
Access Product Name	<string>	Identifies the product the customer currently has purchased (unique within Access Service Provider)
Anonymous Service Instance ID	<string>	Access network SP Anonymous ID provided for correlation and privacy (unique per access service per within an access network SP)

Table 4. [TR-304 Table 4] – Static Access Service Attributes per protocol layer

Attribute	Units	Contents	Description
Protocol		Enumeration of TR-181 interface objects	The protocol layer that the following attributes are for
Provisioned Maximum Down Rate	Bits per second	<integer>	Provisioned limit of downstream rate at this protocol layer
Provisioned Maximum Up Rate	Bits per second	<integer>	Provisioned limit of upstream rate at this protocol layer
Product Minimum Down Rate	Bits per second	<integer>	Per product claim, minimum down rate that the access service can achieve as a

Product Minimum Up Rate	Bits per second	<integer>	general rule at this protocol layer Per product claim, minimum up rate that the access service can achieve as a general rule at this protocol layer
Provisioned Maximum Down Burst	Bits per second	<integer>	Provisioned burst limit of downstream rate at this protocol layer
Provisioned Maximum Up Burst	Bits per second	<integer>	Provisioned burst limit of upstream rate at this protocol layer
Usage limit type	Enumeration of “Unlimited”, “Capped”, “Metered”		Type of usage limit imposed on the access service at this protocol layer. This parameter may only exist or be populated at one of the protocol layers.
Volume cap	MBytes	<integer>	If “Usage limit” = Capped then volume cap per billing cycle at this protocol layer; otherwise null
Throttled Maximum Down Rate	Bits per second	<integer>	If “Usage limit” = Capped and usage is throttled after the cap is exceeded, this represents the limit of downstream rate when the throttle is in place.
Throttled Maximum Up Rate	Bits per second	<integer>	If “Usage limit” = Capped and usage is throttled after the cap is exceeded, this represents the limit of upstream rate when the throttle is in place.

Table 5. [TR-304 Table 5] – Dynamic Access Service Attributes

State	Units	Contents	Description
Current usage	Mbytes	<integer>	Usage for this service, that applies towards a usage limit.
Current datetime	Datetime	The date and time of the Current usage value	Retrieved with Current usage as second field value

Table 8 – Throughput and Rate Metrics

Different measurement methods

Table 9 – Delay Metrics

One-way and two-way, Different measurement methods

Packet Delay Variation (PDV)

Inter-packet and inter-frame delay variation

Loss (PLR)

Continuity = connectivity, UAS

Availability

I.1.1 TR-304 Metric Categories

Categories of metrics in TR-304 are:

Delays

Data rates, throughput, usage

Jitter, Packet Delay Variation (PDV), Inter-packet and inter-frame delay variation

Loss (PLR)

Continuity = connectivity, UAS, Availability

I.2 BBF TR-126

TR-126 [1] describes measurement methods, metrics and targets for those metrics for IPTV, as listed here.

Measurement methods:

- Full Reference (FR) – A method when both the original transmitted and received video signals are available to determine video quality objectively
- Reduced Reference (RR) – A method when partial information about the transmitted video signal and full information about the received video signal are available to determine video quality objectively
- No reference (NR) – A method when only the received video signal is available to determine video quality objectively.

Metrics:

MOS

PSNR

Packet loss, bit rate, latency, jitter

Availability / reliability

Transaction delays, zap time

Resolution, Source quality, encoder and decoder quality

Lip synch

Constituent stream qualities: audio quality, video quality

I.2.1 TR-126 Metric Categories

Categories of metrics in TR-126 are:

MOS

Content quality, source quality/resolution, degradation due to delivery, A/V (e.g., lip synch)

Packet loss, bit rate, latency, jitter

Availability / reliability

Transaction delays, zap time

I.3 BBF TR-452.1, Quality Attenuation Measurement Architecture and Requirements

The TR-452 series of documents define Quality of Experience Delivered (QED), also called Quality Attenuation (written ΔQ). QED is an overarching metric class that gives a complete characterization of packet performance including delay and loss experienced across network segments and end-to-end. TR-452.1 [6] defines a reference architecture and specifies requirements for measuring and analyzing quality attenuation on paths and sub-paths of a broadband network. The theory, principles, use cases and measurement approach are also presented.

I.4 BBF TR-421, Application-Layer Test Traffic Architecture and Requirements

TR-421 [5] defines an architecture and related requirements for the specification of test traffic and measurements associated with the application layer. This gives an overview of the approximate characteristics (but not actual test streams) for application classes: streaming video, VoIP, video conferencing, web browsing, security video, file transfer, peer-to-peer file sharing, gaming, AR/VR, and performance testing.

I.5 BBF-TR-471, Maximum IP-Layer Capacity Metric, Related Metrics, and Measurements

TR-471 [9] defines UDP-based speed tests (UDPST) that accurately and efficiently measure broadband bit rate and latency. UDPST can easily generate stream patterns with sending rates that vary according to a time schedule, to mimic different traffic types. UDPST can also use its existing feedback measurements of loss and delay to add dynamic adaptation to application streams, in order to follow a specific algorithm (many applications have some rate-adaptation to their connection conditions).

I.6 3GPP TS 28.554 KPIs

3GPP TS 28.554 [15] specifies end-to-end Key Performance Indicators (KPIs) for the 5G network and network slicing. An overall listing of the KPIs in TS 28.554 are listed here. Note that additional 3GPP references provide information related to metrics [10][11][12][13][14].

6.2 Accessibility KPI

6.2.1 Mean registered subscribers of network and network slice through AMF (Access and Mobility Management Function)

6.2.2 Registered subscribers of network through UDM

6.2.3 Registration success rate of one single network slice

6.2.4 Partial DRB (Data Radio Bearer) Accessibility for UE services

6.2.5 PDU session Establishment success rate of one network slice (S-NSSAI)

6.2.6 Maximum registered subscribers of network slice through AMF

- 6.2.7 Total DRB accessibility for UE services
- 6.2.8 Mean CM (Connection Management) Connected subscribers of network slice through AMF
- 6.2.9 Maximum on-line subscribers of network slice through AMF
- 6.2.10 PFCP session established success rate of one network and one network slice
- 6.3 Integrity KPI
 - 6.3.1 Latency and delay of 5G networks
 - 6.3.1.1 Downlink latency in gNB-DU
 - 6.3.1.2 Integrated downlink delay in RAN
 - 6.3.1.3 Downlink delay in gNB-DU
 - 6.3.1.4 Downlink delay in gNB-CU-UP
 - 6.3.1.5 Uplink delay in gNB-DU
 - 6.3.1.6 Uplink delay in gNB-CU-UP
 - 6.3.1.7 Integrated uplink delay in RAN
 - 6.3.1.8 E2E delay for network slice
 - 6.3.2 Upstream throughput for network and Network Slice Instance
 - 6.3.3 Downstream throughput for Single Network Slice Instance
 - 6.3.4 Upstream Throughput at N3 interface
 - 6.3.5 Downstream Throughput at N3 interface
 - 6.3.6 RAN UE Throughput
 - 6.3.6.2 RAN UE Throughput definition
 - 6.3.6.3 DL RAN UE throughput
 - 6.3.6.4 UL RAN UE throughput
- 6.4 Utilization KPI
 - 6.4.1 Mean number of PDU sessions of network and network Slice Instance
 - 6.4.2 Virtualized Resource Utilization of Network Slice Instance
 - 6.4.3 PDU session establishment time of network slice
 - 6.4.4 Mean number of successful periodic registration updates of Single Network Slice
 - 6.4.5 Maximum number of PDU sessions of network slice
- 6.5 Retainability KPI
 - 6.5.1 QoS flow Retainability (how often an end-user abnormally loses a QoS flow)
 - 6.5.2 DRB Retainability
- 6.6 Mobility KPI
 - 6.6.1 NG-RAN handover success rate
 - 6.6.2 Mean Time of Inter-gNB handover Execution of Network Slice
 - 6.6.3 Successful rate of mobility registration updates of Single Network Slice
 - 6.6.4 5GS to EPS handover success rate
 - 6.6.5 NG-RAN handover success rate for all handover types

- 6.7 Energy Efficiency (EE) KPI
 - 6.7.1 NG-RAN data Energy Efficiency (EE)
 - 6.7.2 Network slice Energy Efficiency (EE)
 - 6.7.2.1 Generic Network Slice Energy Efficiency (EE) KPI
 - 6.7.2.2 Energy efficiency of eMBB network slice
 - 6.7.2.2a Energy efficiency of eMBB network slice – RAN-based
 - 6.7.2.3 Energy efficiency of URLLC network slice
 - 6.7.2.4 Energy efficiency of MIoT network slice
 - 6.7.3 5G Energy Consumption (EC)
 - 6.7.3.1 NF Energy Consumption (EC)
 - 6.7.3.1.2 Estimated Virtualized Network Function (VNF) energy consumption
 - 6.7.3.1.3 Estimated Virtualized Network Function Component (VNFC) energy consumption
 - 6.7.3.1.4 Estimated virtual compute resource instance energy consumption based on mean vCPU usage
 - 6.7.3.2 5GC Energy Consumption (EC)
 - 6.7.3.3 Network Slice Energy Consumption (EC)
 - 6.7.3.4 NG-RAN Energy Consumption (EC)
 - 6.7.3.4.1 NG-RAN EC
 - 6.7.3.4.2 gNB EC
 - 6.7.4 5GC Energy Efficiency (EE)
 - 6.7.4.1 Generic 5GC Energy Efficiency (EE) KPI
 - 6.7.4.2 Energy Efficiency of 5GC based on the useful output of 5GC user plane
- 6.8 Reliability KPI
 - 6.8.1.1 Packet transmission reliability KPI in DL on Uu
 - 6.8.1.2 Packet transmission reliability KPI in UL on Uu
 - 6.8.1.3 Packet transmission reliability KPI in DL on N3
 - 6.8.1.4 Packet transmission reliability KPI in UL on N3

I.6.1 3GPP TS 28.554 Metric Categories

Categories of metrics in TS 28.554 are:.

Throughput

Latency

Accessibility

Registered number of subscribers (pre slice, total)

Success rate

Utilization

Number of PDU sessions
 Resource utilization [exhaustion]
 Successful registrations

Retainability

Retainability (how often an end-user abnormally loses a QoS flow)
 Handover

Energy

Energy efficiency
 Energy consumption

Reliability**Packet transmission reliability****Mobility**

N/A

I.7 ATIS IPTV Interoperability Forum (IIF)

The ATIS IPTV Interoperability Forum (IIF) QoS Metrics committee defined a number of standards with metrics for IPTV. Pertinent IPTV metric standards are given here along with lists of the metrics in these. ATIS IIF QoS metric standards of interest are referenced here.

I.7.1 ATIS-0800008.v002, Quality of Service Metrics for Linear (Broadcast) IPTV

ATIS-0800008.v002 [38] defined QoS Metrics for linear (broadcast) IPTV which include:

Transaction Quality, Content Quality, Media Stream Quality, Transmission Quality.

Domains: Content provider, Service provider, Network provider, Consumer.

Frame loss rates (before and after correction), burst loss characteristics (bursts, gaps), out-of-order packets, discards, etc.

MPEG TS stats, per PID

Audio dropouts, audio loudness, AV synch

Measurement interval

MOS; min, max, algorithm. Audio, video

Advantage factor: This factor represents the customer's acceptance of some decrease in quality for access advantage, e.g., mobility or connection into hard-to-reach regions.

Jitter, Peak Packet-to-Packet Delay Variation (PPDV)

Compression characterization parameters

Buffer underruns and overruns

Content quality

Multicast delays (zap times), STB boot time, service boot time

Availability

Program guide
VoD, transaction times,
Advertising
Fault codes

I.7.2 ATIS-0800041, Implementers Guide to IPTV QoS Metrics

ATIS-0800041 [41] further presents metrics. Some salient tables of interest for content quality, alarm levels, and error characterization are shown here.

Table 6. [ATIS-0800041] Quality and Impairment scales, Five Grade Scale
(Source: ITU-R BT.500)

Quality (MOS)		Impairment	
5	Excellent	5	Imperceptible
4	Good	4	Perceptible, but not annoying
3	Fair	3	Slightly annoying
2	Poor	2	Annoying
1	Bad	1	Very annoying

A stratification of alarm levels typically used by the industry is presented in Table 7. These types of alarm classifications are common for monitoring physical layer links, interfaces, and equipment. Note that there is often only one level for critical and severe levels, not both.

Table 7. [ATIS-0800041] Alarm Levels

Critical	1	Worst
Severe	2	
Major	3	
Minor	4	
No trouble	5	Best

Table 6 and Table 7 show numerical scales from 1 (worst) to 5 (best). These roughly correspond between the tables.

The error characterizations in Table 8 were originally developed and used for MPEG transport streams [39].

Table 8. [ATIS-0800041] Error Characterization

Transport Stream Off Air (TOA)	1	Worst
Component Missing (CM)	2	
Quality Of Service (QOS)	3	
Technically Non-Conformant (TNC)	4	
No Error	5	Best

I.7.3 ATIS IIF Metric Categories

Categories of metrics defined by ATIS IIF are:

Transaction Quality, Content Quality, Media Stream Quality, Transmission Quality

Domains: Content provider, Service provider, Network provider, Consumer

Loss rates, dropouts.

Advantage factor: This factor represents the customer's acceptance of some decrease in quality for access advantage, e.g., mobility or connection into hard-to-reach regions.

Jitter, Peak Packet-to-Packet Delay Variation (PPDV). Out of order (OOO) packets.

Compression characterization parameters

Buffer underruns and overruns

Content quality

Multicast delays (zap times), STB boot time, service boot time, VoD, transaction times,

Availability

Program guide

Advertising

Fault codes

I.8 ETSI GS NFV-TST 008

ETSI GS NFV-TST 008 [42] defines metrics for the NFV Infrastructure, which include:

Compute Metrics

Processor Usage

Processor Utilization

Network Metrics

Packet count.

Octet count.

Dropped Packet count.

Errored Packet count.

Memory Metrics

Memory buffered

Memory Cached

Memory free

Memory Slab

Memory Total

Memory Used

memory page size metrics

Hypervisor/MANO usage

Disk, GPU

I.9 5G-PPP White Paper: Beyond 5G/6G KPIs and Target Values

5G-PPP White Paper: Beyond 5G/6G KPIs and Target Values [43] describes KPIs and their target values for 6G, which are:

2 Standard Network KPI Definitions

- 2.1 Peak Data Rate
- 2.2 User Experienced Data Rate
- 2.3 Area Traffic Capacity
- 2.4 Bandwidth
- 2.5 Connection Density
- 2.6 Latency
 - 2.6.1 User Plane Latency
 - 2.6.2 Control Plane Latency
- 2.7 Reliability
- 2.8 Peak Spectral Efficiency
- 2.9 5th Percentile User Spectral Efficiency
- 2.10 Average Spectral Efficiency
- 2.11 Energy Efficiency
- 2.12 Energy Efficiency in NFV
- 2.13 Higher-Accuracy Positioning
- 2.14 Quality of Experience

3 Clustering of KPIs Provided by ICT-52 Projects

- 3.1 Latency
- 3.2 Capacity
- 3.3 Packet Loss
- 3.4 Compute
- 3.5 Energy
- 3.6 Security
- 3.7 Channel
- 3.8 Electromagnetic Field (EMF)
- 3.9 Localization
- 3.10 Service Availability and Reliability

I.10 Cloud Service Metrics

Cloud providers and hyperscalers including Amazon Web Services [75], Alphabet, Meta, Azure, etc. define metrics for cloud service, including:

Uptime, availability
Mean time between failure (MTBF)
Mean time to repair (MTTR)
CPU, Fast path CPU
Memory, volatile, non-volatile
Queue monitoring
Disk
Requests per minute
Time to acknowledge or respond
Latency
Error rate
Swap usage
Security
Throughput (transactions per second)
Capacity, number of users, sessions, CPU, memory, VMs/containers, etc.
Scalability
Load Balancing
Feature release time

I.11 IETF IPPM

IETF IP Performance Metrics (IPPM) has defined many metrics and metric measurement methods in detail. Internet drafts and RFCs in IPPM are [18][20][23][28][30][32][33][34]. IPPM has very detailed and unambiguous standard metric definitions.

Measurement protocols defined by IETF include:

- RFC 4656 A One-way Active Measurement Protocol (OWAMP) [25]
- RFC 5357 A Two-Way Active Measurement Protocol (TWAMP) [27]
- RFC 8762 Simple Two-Way Active Measurement Protocol (STAMP) [35]

Also, Precision Availability Metrics (PAM) [37] are aimed at capturing end-to-end service levels for a flow; as well as standards for metrics for capacity, rate measurement, delay, packet loss, loss patterns, loss episodes, packet delay variation, connectivity monitoring, flow measurement, metric composition, metric registry and in-situ measurements and OAM.

I.12 IANA Performance Metrics Registry

The IANA Performance Metrics Registry is posted online [17]. This has many network-level metrics.

I.13 Metro Ethernet

Metro Ethernet metrics and measurement architectures have been defined by the Metro Ethernet Forum (MEF), ITU-T and IEEE. These include Layer 2 Ethernet Connectivity Fault Management (CFM); Continuity Check Protocol (CCP, including heartbeat), loopback (layer 2 ping), linktrace (layer 2 traceroute) and alarming.

ITU-T G.8013/Y.1731 and IEEE 802.1ag [68] define operation, administration, and maintenance (OAM) functions and mechanisms for Ethernet-based networks. This includes Ethernet continuity check (ETH-CC), Ethernet loopback (ETH-LB), Ethernet link trace (ETH-LT), Ethernet alarm indication signal (ETH-AIS), Ethernet remote defect indication (ETH-RDI, used with ETH-CC), Ethernet test signal (ETH-Test, for testing one-way bandwidth, loss), and Ethernet automatic protection switching (ETH-APS). Section 8 defines Ethernet frame loss measurement (ETH-LM), Frame delay measurement (ETH-DM), Throughput measurement, and Synthetic loss measurement (ETH-SLM), including detailed specification with a number of options.

MEF 38, Service OAM Fault Monitoring YANG Module [72], defines metrics for topology, continuity check, loopback, linktrace, Alarm indication signal, test signal.

MEF 39, Service OAM Performance Monitoring YANG Module [73], defines metrics for topology, continuity check, loopback, linktrace, Alarm indication signal, test signal.

ITU-T G.8052.1 [63] defines Ethernet OAM in the ITU.

I.14 IEEE Std 802.3.2, IEEE Standard for Ethernet YANG Data Model Definitions

IEEE Std 802.3.2 [44] defines YANG modules for power delivery, EPON (including transmit power, FEC counters, queues, and multicast), and Ethernet Link OAM (ELM) (including remote fault indication, link monitoring, remote loopback, and frame counters).

End of Broadband Forum Technical Report TR-499