DSL Forum

Technical Report
TR-126

Triple-play Services
Quality of Experience (QoE)
Requirements

13 December 2006

Produced by:
Architecture & Transport Working Group

Co-editors:
Tim Rahrer, Nortel
Riccardo Fiandra, FastWeb
Steven Wright, BellSouth

Working Group Co-chairs:
David Allan, Nortel
David Thorne, BT

Notice

The DSL Forum is a non-profit corporation organized to create guidelines for DSL network system development and deployment. This Technical Report has been approved by members of the Forum. This document is not binding on the DSL Forum, any of its members, or any developer or service provider. This document is subject to change, but only with approval of members of the Forum.

This publication may incorporate intellectual property. The DSL Forum encourages but does not require declaration of such intellectual property. For a list of declarations made by DSL Forum member companies, please see www.dslforum.org.

Notwithstanding anything to the contrary, the DSL Forum makes no representation or warranty, expressed or implied, concerning this publication, its contents or the completeness, accuracy, or applicability of any information contained in this publication. No liability of any kind shall be assumed by the DSL Forum as a result of reliance upon any information contained in this publication. The DSL Forum does not assume any responsibility to update or correct any information in this publication.

Document History

Revision History	Date	Version Editor	Reason for Update
Version 0.1	February 22, 2005	Tim Rahrer - Nortel	Initial skeleton draft based on contribution 2004.507.01 and 2004.513.01 presented at the Orlando meeting, December 2004.
Version 0.2	April 27, 2005	Tim Rahrer - Nortel	Update from initial draft and includes expanded scope to triple-play services agreed to in Q1 Austin meeting, changes recommended in Austin, additional details on video quality measurement and first draft of video control plane requirements and transport layer QoE requirements for video for discussion.
Version 0.3	September 12, 2005	Tim Rahrer - Nortel	Corrections and fixes. Standard definition application layer performance guidelines added. Interactive (gaming) applications moved to a future work section due to lack of inputs.

Version 0.4	November 29, 2005	Tim Rahrer - Nortel, Riccardo Fiandra - FastWeb	Added inputs application and transport layer inputs from Q3 DSLF meeting (dsl2005.528) and updated HD performance targets. Discussion of FEC and related materials aimed at chapter 7, (dsl2005.528 and dsl2005.460) require further study and analysis so are delayed until the next update of WT-126.
Version 0.5	February 21, 2006	Tim Rahrer - Nortel	Draft Voice application QoE requirements added. Minor edits and corrections.
Version 0.6	April 24, 2006	Tim Rahrer - Nortel, Riccardo Fiandra - FastWeb	Changed IP loss period to 5 IP packets max to better align with practicalities of xDSL impairments and FEC mechanisms. Added draft section on subjective results.
Version 0.7	May 2, 2006	Tim Rahrer - Nortel, Riccardo Fiandra – FastWeb, Steven Wright - BellSouth	Incorporated changes recommended on conf. call including more clearly defining scope to QoE for triple play services. Added dependability factors. Reworked subjective quality experiment data into an appendix
Version 0.8	July 2006	Tim Rahrer - Nortel	Added final comments from San Jose inputs. Additional edits and clean ups also done. Significant additions to Best-effort data section.
Version 0.85	July 2006	Tim Rahrer – Nortel	Added inputs on packet loss period changing from fixed number of packets to time duration (dsl2006.541). Added additional text on gaming requirements (dsl2006.544).
Version 0.9	July 2006	Tim Rahrer - Nortel, Riccardo Fiandra – FastWeb, Steven Wright - BellSouth	Updated Appendices
Version 1.06	Sept. 17, 2006	Tim Rahrer, Nortel	Integrated Straw Ballot comments and contributions in advance of Q3 meeting.
Version 1.1	Sept. 25, 2006	Tim Rahrer, Nortel	Integration of straw ballot review comments from Athens meeting
Version 1.2	October 25, 2006	Tim Rahrer, Nortel	Completed Straw Ballot comment and inputs. Updated list of acronyms
Version 1.3	November 2, 2006	Tim Rahrer, Nortel	Minor typos and edits for clarification before Letter Ballot.
Version TR	December 13, 2006	Tim Rahrer, Nortel	Update to text to reflect TR status.

Technical comments or questions about this document SHOULD be directed to:

Editors: Tim Rahrer
Nortel
3500 Carling Ave., Ottawa, ON Canada
Tel : +1.613.763.4479
Email : tim.rahrer@nortel.com

Riccardo Fiandra
FastWeb, SpA
Email : riccardo.fiandra@fastweb.it

Steven Wright
BellSouth Telecommunications
Email : steven.wright@bellsouth.com

Table of Contents

List of Figures

List of Tables

SUMMARY

This Technical Report presents the recommended minimum end-to-end quality of experience (QoE) guidelines in terms of engineering objective measures for triple play applications delivered through a broadband infrastructure. Some informative implementation alternatives to meet the QoE targets are also discussed in Appendices. QoE requirements define the overall, performance at the services level from the perspective of the end user. The establishment of consistent, baseline subjective QoE for end users and corresponding objective engineering targets is critical to the market success of broadband service offerings. The QoE guidelines presented in this document are end-to-end requirements and are agnostic to access technology (xDSL, xPON, etc.), services architecture, and implementation. Initial applications presented are entertainment video (video on demand, and broadcast video), voice, and best-effort data (web browsing, gaming). Other applications such as, video conferencing may be included in a future revision.

DSL Forum Technical Report
TR-126

Triple-play Services Quality of Experience (QoE) Requirements

1. Purpose

The purpose of this Technical Report is to present the recommended minimum end-to-end quality of experience (QoE) requirements in terms of objective engineering measures for triple play applications delivered through a broadband infrastructure. Some informative implementation alternatives to meet the QoE targets are also discussed in the Appendices.

QoE requirements define the overall, subjective performance at the services level from the perspective of the end user. The establishment of consistent, baseline quality of experience for end users and corresponding objective engineering targets is critical to the market success of broadband service offerings. The QoE guidelines and objective performance recommendations presented in this document are in the context of an end-to-end system and are agnostic to access technology (xDSL, xPON, etc.), services architecture, and implementation. Initial applications presented are entertainment video (video on demand and broadcast video), voice, and best-effort data (e.g. web browsing). Other applications such as video conferencing may be included in a future revision.

2. Scope

The primary scope of this document is the definition of user requirements (Quality of Experience – QoE) in the form of objective engineering targets in the support of key triple play services over a broadband access architecture. These new services can be offered to residential mass market users as well as to small and medium sized enterprises (SMEs). The Quality of Experience requirements are defined from an end user perspective. The QoE requirements are agnostic to network deployment architectures and transport protocols (ex. IP or ATM), access network technology (xDSL, xPON, wireless, etc.), and are meant to set the minimum requirements for a satisfactory user experience.

The QoE requirements are specified as end-to-end (not just access link) and can be translated into objective engineering measures at the network transport and application layers given various assumptions concerning the network and service architectures. Service providers should ensure that the minimum objective engineering performance recommendations defined in this document are met but may wish to set their own preferred and/or premium QoE driven targets to provide differentiated services in their markets. Additionally an informational section on factors influencing architecture alternative decisions, and an overview of loss mitigation mechanisms is presented. References are made to related work in other organizations.

The QoE requirements put forward in this document focus on key broadband entertainment and communications services to end users. These services are:

- Video on Demand (VoD) to fixed TV viewing devices
- Broadcast Video to fixed TV viewing devices
- Voice Over IP
- Best-effort (BE) Internet access (e.g. web browsing and gaming).

Video, whether broadcast or on demand, voice and best-effort Internet access together are referred to as triple-play services.

Others applications and services that may be addressed in a future version include:

- Interactive multimedia (animation, music on demand, web camera control, etc.)
- Video conferencing
- Secure data (VPN)
- FTP file transfer
- P2P
- Streaming video to a PC.

For the purposes of this document, professional level video networks such as studio content to edit facility, live event backhaul, contribution networks for rebroadcast are out of scope and left to other forums such as ProMPEG[16], Video Services Forum (VSF)[18], and ATIS Network Performance and QoS Subcommittee (NPQSC)[21].

The success of triple-play services requires that they meet user needs and expectations and deliver a satisfactory quality of experience (QoE). Additionally, services provided over a broadband infrastructure must perform at least as well as, and preferably superior to competing services using other delivery mechanisms. Consumers tend to be agnostic to broadband transport mechanisms whether they be xDSL, fiber, wireless, cable, etc. as long as the infrastructure does not get in the way of their goals. Analyzing service QoE requirements (end-to-end) and establishing corresponding objective engineering performance targets is the first step to ensuring end users are satisfied with their services.

While general implementation factors that impact service performance and an overview on loss mitigation mechanisms will be provided, deployment architecture and implementation decisions are left to other working groups and the service providers themselves. This document provides a stake in the ground for making such decisions in terms of the user requirements. Other details such as cost, existing infrastructure, network evolution, subscriber numbers and scalability, supplier choice, manageability, content aggregation scalability, etc. will need to be factored as well in future work by standards bodies and/or service providers.

2.1 Definitions

The following definitions apply for the purposes of this document:

Application Layer: - the level where various application parameters are set, for example media resolution, encoder type, bit rate, decoder loss concealment, etc. May also include application layer forward error correction.

CBR Traffic: - constant bit rate – refers in this document to the bit rate of the video stream at the output of the encoder (application layer) where the traffic pattern information units arrive at a constant bit rate measured over a specified time interval and not to the ATM transport concepts of CBR.

End-to-end: - refers to the complete system from source of an application (media server, satellite uplink, application server, etc.), through the all network segments (national, regional, access) to the end user customer premises network and devices (TV, computer, etc.) used to view / consume the application or service. For the purposes of this document, end-to-end includes any application layer error correction mechanisms that may reside in consumer premises equipment.

MOS - Mean Opinion Score: - The MOS is generated by averaging the results of a set of standard, subjective tests where a number of users rate the quality on a five point scale from 1 (Bad / Very Annoying) to 5 (Excellent / Imperceptible impairments). The MOS is the arithmetic mean of all the individual scores.

PESQ - Perceptual Evaluation of Speech Quality: Is an ITU standard (P.862) for measuring speech quality.

Quality of Experience (QoE): - is the overall performance of a system from the point of view of the users. QoE is a measure of end-to-end performance at the *services level* from the *user perspective* and an indication of how well the system meets the user's needs.

Quality of Service (QoS): - is a measure of performance at the *packet leve*l from the *network perspective*. Quality of Service (QoS) also refers to a set of technologies (QoS mechanisms) that enable the network operator to manage the effects of congestion on application performance as well as providing differentiated service to selected network traffic flows or to selected users

Service Layer: - the layer exposed to the user, where QoE is measured

Transport Layer: - it is the layer responsible for transporting service packets from one entity to another. It employs routing and forwarding to achieve its function. It may also employ traffic management mechanisms as needed. Various network impairments (loss, delay, jitter) can occur here and is also where QoS and error correction mechanisms may be employed. For the purposes of this document includes Layer 4 and below from OSI networking model.

VBR: - variable bit rate – refers in this document to the bit rate of the video stream at the output of the encoder (application layer) and not to the ATM transport concepts of VBR.

2.2 Abbreviations

The following abbreviations apply for the purposes of this document:

AAA	Authentication, Authorization, and Accounting
AAL5	ATM Adaptation Layer 5
ACK	Acknowledgement
ADSL	Asymmetric Digital Subscriber Line
ADS	Asymmetric Digital Subscriber Line
ARP	Address Resolution Protocol
ARQ	Automatic Repeat Request
ASP	Application Service Provider
ATM	Asynchronous Transfer Mode
AVC	Advanced Video Coding, an MPEG-4 profile (also known as H.264 or MPEG-4, Part 10)
B-NT	Broadband Network Termination
BB	Broadband
BE	Best-effort
BER	Bit error rate
BoD	Bandwidth on Demand
BRAS	Broadband Remote Access Server
CBR	Constant Bit Rate
CATV	Cable TV
CO	Central Office
CPE	Customer Premises Equipment
CPN	Customer Premises Network
DCT	Discrete Cosine Transform
DHCP	Dynamic Host Configuration Protocol
DiffServ	Differentiated Services
DRM	Digital Rights Management
DSL	Digital Subscriber Line
DSLAM	Digital Subscriber Line Access Multiplexer
DVR	Digital Video Recorder
EPG	Electronic Program Guide
FEC	Forward Error Correction
FPS	First Person Shooter (game)
FTP	File Transfer Protocol
HD	High Definition
IEEE	Institute of Electrical and Electronics Engineers
IETF	Internet Engineering Task Force
IGMP	Internet Group Management Protocol
IM	Instant Messaging
IP	Internet Protocol
ISP	Internet Service Provider

LAN	Local Area Network
LLC	Logical Link Control
MAC	Medium Access Control
MDU	Multi-user Dungeon (gaming)
MMO	Massively Multiplayer Online (gaming)
MMOFPS	Massively Multiplayer Online First Person Shooter (game)
MMORPG	Massively Multiplayer Online Role Playing Game
MMORTS	Massively Multiplayer Online Real Time Strategy (game)
MOS	Mean Opinion Score
MPEG	Motion Pictures Expert Group
MPLS	Multi-Protocol Label Switching
MSO	Multi-service Operator
MSS	Maximum Segment Size
MTBE	Mean Time Between Errors
MTU	Message Transfer Unit
NAT	Network Address Translation
NAPT	Network Address and Port Translation
NSP	Network Service Provider
P2P	Peer to peer
PC	Personal Computer
PESQ	Perceptual Evaluation of Speech Quality
PLR	Packet Loss Rate
PON	Passive Optical Network
POTS	Plain Old Telephone Service
PPP	Point-to-Point Protocol
PPPoE	Point-to-Point Protocol over Ethernet
PPV	Pay per view
PS	POTS Splitter
PVC	Permanent Virtual Circuit
PVR	Personal Video Recorder
QoE	Quality of Experience
QoS	Quality of Service
RADIUS	Remote Authentication Dial-In User Service
RFC	Request For Comments
RF	Radio Frequency
RG	Routing Gateway / Residential Gateway
RPG	Role Playing Game
RTS	Real Time Strategy (gaming)
RTT	Round Trip Time
SD	Standard Definition
SIP	Session Initiation Protocol
SLA	Service Level Agreement
SM	Service Module
SME	Small to Medium Sized Enterprise
SONET	Synchronous Optical Network
SRT	System Response Time

SSM	Source Specific Multicast
STB	Set-Top Box
SVC	Switched Virtual Circuit
TCP	Transmission Control Protocol
TV	Television
UDP	User Datagram Protocol
VBR	Variable Bit Rate
VC	Virtual Circuit
VLAN	Virtual Local Area Network
VoD	Video on Demand
VoDSL	Voice over DSL (could be IP or ATM-based)
VoIP	Voice over Internet Protocol
VPN	Virtual Private Network
VQEG	Video Quality Expert Group
xTU-C	xDSL Termination Unit - Central Office
xTU-R	xDSL Termination Unit - Remote

2.3 Conventions

In this document, several words are used to signify the requirements of the specification. These words are often capitalized.

MUST	This word, or the adjective "REQUIRED", means that the definition is an absolute requirement of the specification.
MUST NOT	This phrase means that the definition is an absolute prohibition of the specification.
SHOULD	This word, or the adjective "RECOMMENDED", means that there may exist valid reasons in particular circumstances to ignore this item, but the full implications must be understood and carefully weighted before choosing a different course.
MAY	This word, or the adjective "OPTIONAL", means that this item is one of an allowed set of alternatives. An implementation that does not include this option MUST be prepared to inter-operate with another implementation that does include the option.

3. References

No normative references are included in this document.

The following information is given for the convenience of users of this TR and does not constitute an endorsement by the DSL Forum of these products. Equivalent products MAY be used if they can be shown to lead to the same results.

Adobe® is either registered trademarks or trademarks of Adobe Systems Incorporated in the United States and/or other countries.
CableLabs® is a registered trademark of Cable Television Laboratories, Inc.
HomePlug® is a registered trademark of HomePlug Powerline Alliance, Inc.
HomePNA® is a registered trademark of HomePNA, Inc.
Microsoft® is a registered trademark of Microsoft Corporation.
PowerPoint® is a registered trademark of Microsoft Corporation.
PESQ® is a registered trademark of Psytechnics Limited.
Telcordia® is a registered trademark of Telcordia Technologies, Inc.

3.1 Relation of this document to other documents

One goal of this Technical Report is to provide a foundation for other work groups discussing architecture alternatives, QoS mechanisms, etc. Establishment of minimum service requirements from an end user perspective will ensure that the evolution of broadband access (e.g. DSL, PON, etc.) technology will support the expectations of service consumers and enable the continued success of broadband access technology and broadband access globally.

The requirements presented in this document refer to and/or can be used by other DSL Forum Technical Reports, Working Texts and Proposed Drafts, specifically those targeted to architectures and QoS mechanisms. These include (but are not limited to) the following:

- TR-058: Multi-Service Architecture & Framework Requirements[1];
- TR-059: DSL Evolution - Architecture Requirements for the Support of QoS-Enabled IP Services[2]
- TR-068: Routing gateway modem requirements[3];
- TR-069: Auto-configuration of advanced services[4];
- TR-094: Multi-Service Delivery Framework for Home Networks[5]
- TR-101: Migration to Ethernet Based DSL Aggregation[6]
- TR-102: Service Interface Requirements for TR-058 Architectures[7]

Where possible / practical, application performance parameters are based on other related industry forums and standards bodies including (non-exhaustive list):

- ITU-T G.1000: Communications quality of service: A framework and definitions[8]
- ITU-T G.1010: End-user multimedia QoS categories[9]
- ITU-T G.1020: Performance parameter definitions for quality of speech and other voiceband applications utilizing IP networks[10]
- ITU-T G.107: The E-model, a computational model for use in transmission planning.[11]
- ITU-T Y.1541: Network performance objectives for IP-based services[12] (including proposed updates)
- ITU-T J.241: Quality of Service ranking and measurement methods for digital video services delivered over broadband IP Networks[13]

- ITU-T J.series on video quality measurement including J.140-149, [14]
- VQEG: Video Quality Experts Group[15]
- ProMPEG: Professional MPEG Forum[16]
- DLNA Digital Living Network Alliance[17]
- VSF Video Services Forum[18]
- CableLabs®: CableLabs® Video-On-Demand Content Encoding Profiles Specification[19]
- IETF: Internet Engineering Task Force, IP Performance Metrics (IPPM) Charter[20]
- ATIS: Network Performance and QoS Subcommittee (NPQSC) (formerly T1A1.3)[21]
- ATIS: IPTV Interoperability Forum (IIF), particularly the Architecture and QoS Metrics (QoSm) task forces.[22]
- ETSI: Speech Processing, Transmission and Quality Aspects (ETSI STQ)[23]
- ETSI: TISPAN, particularly Working Group 1[24]
- DVB: DVB-IP Project Phases 1.x and 2 on IPTV delivery of DVB services[25]

The QoE target guidelines presented in this document can be used by other groups when studying the tradeoffs between various architectures, QoS mechanisms and implementations. The remainder of the document provides details for each application in the following categories:

- End-to-end QoE dimensions
- QoE measurement
- QoE targets.

The document closes with an introduction to loss protection mechanisms in the Appendices that may be employed to achieve the QoE targets on the critical dimension of packet loss.

4. Introduction

This section provides an introduction to quality of experience (QoE), distinction between QoE and QoS as used in this document, dimensions of QoE and QoE engineering and the concepts used throughout the remainder of the document. Details of the applications addressed by this document are also provided.

4.1 Quality of Experience (QoE) and Quality of Service (QoS)

QoE (quality of experience) and QoS (quality of service) terminology are often used interchangeably but are actually two separate concepts. QoE is the overall performance of a system from the point of view of the users. QoE is a measure of end-to-end performance at the services level from the user perspective and an indication of how well the system meets the user’s needs. One QoE measurement metric is the Mean Opinion Score (MOS). Mean Opinion Scores are typically used as a subjective measurement to quantify the perceptual impact (the users’ quality of experience) of various forms of service degradation. Other objective metrics of service quality such as the duration of periods of degraded service (e.g. Degraded Seconds, Errored Seconds, Unavailable Seconds) provide less information but may be easier to measure. QoE is also studied and formally defined in the ITU-T Study Group 12[26], specifically in Question 13/12[27] and in the ATIS IPTV Interoperability Forum (IIF), QoS Metrics Task Force (QoSM)[22].

QoS is a measure of performance at the packet level from the network perspective. QoS also refers to a set of technologies (QoS mechanisms) that enable the network administrator to manage the effects of congestion on application performance as well as providing differentiated service to selected network traffic flows or to selected users. QoS metrics may include network layer measurements such as packet loss, delay or jitter.

In general there is a non linear relationship between the subjective QoE as measured by the MOS and various objective parameters of service performance (e.g. encoding bit rate, packet loss, delay, availability, etc.) as shown in Figure 1.

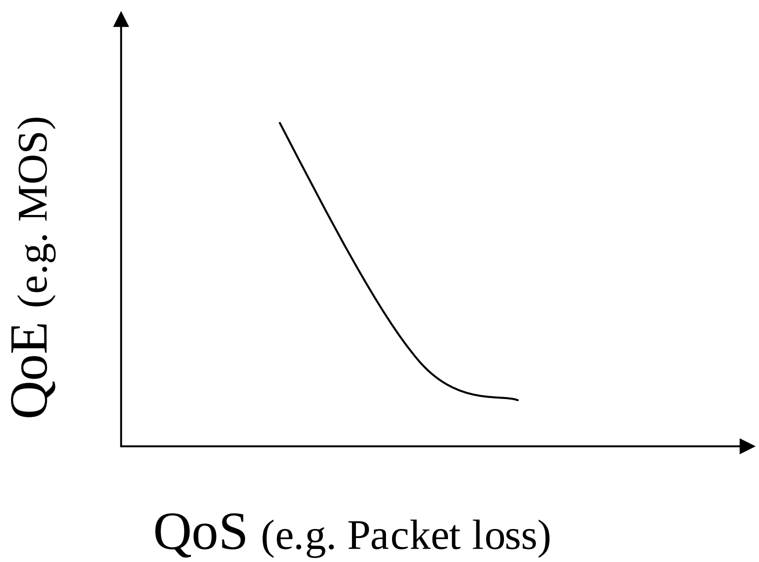


Figure 1 QoE relationship to QoS

Typically there will be multiple service level performance (QoS) metrics that impact overall QoE. The relation between QoE and service performance (QoS) metrics is typically derived empirically. Having identified the QoE/QoS relationship, it can be used in two ways:

a. Given a QoS measurement, one could predict the expected QoE for a user
b. Given a target QoE for a user, one could deduce the net required service layer performance.

To ensure that the appropriate service quality is delivered, QoE targets should be established for each service and be included early on in system design and engineering processes where they are translated into objective service level performance metrics. Quality of experience will be an important factor in the marketplace success of triple-play services and is expected to be a key differentiator with respect to competing service offerings. Subscribers to network services do not care how service quality is achieved. What matters to them is how well a service meets their expectations for effectiveness operability, availability, and ease of use.

4.1.1 End-to-end view

QoE requirements and service provider networks must be considered in a complete end-to-end system. All application services and network elements that can contribute to consumer experience in using a service must be accounted for including:

- Application service nodes (example video head end)
- National and regional networks
- Access network
- Home network
- Customer premises equipment
- Application terminal (example STB and TV).

Figure 2 illustrates a high level end-to-end services architecture as defined in TR-059[2] Figure 3 shows the Ethernet-based equivalent as defined in TR-101[6]. Note that QoE performance targets are specified with respect to the user and are independent of network transport mechanisms such as ATM-based in TR-059 or Ethernet-based links in TR-101.

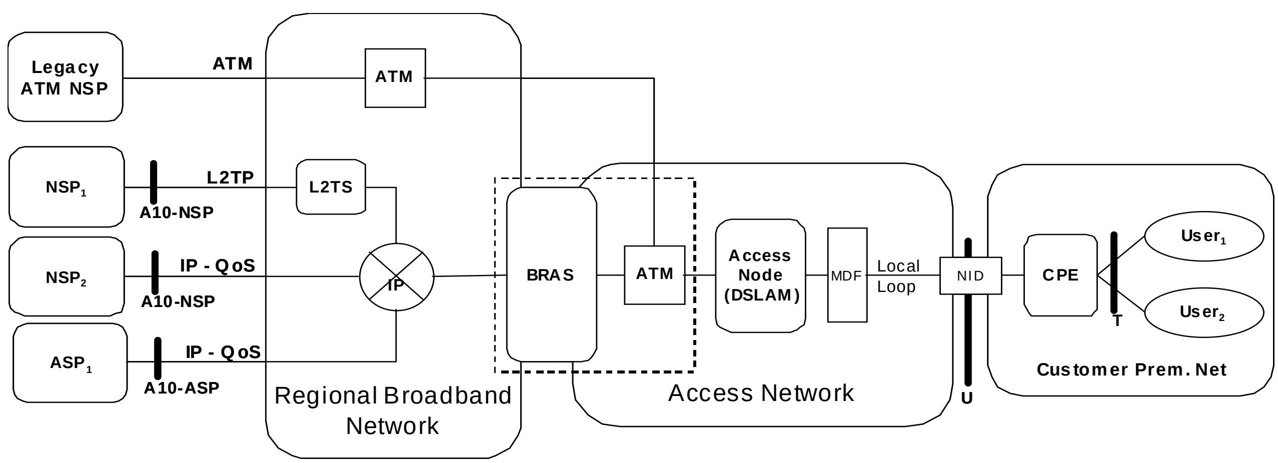


Figure 2 TR-059 High Level Reference Architecture Example

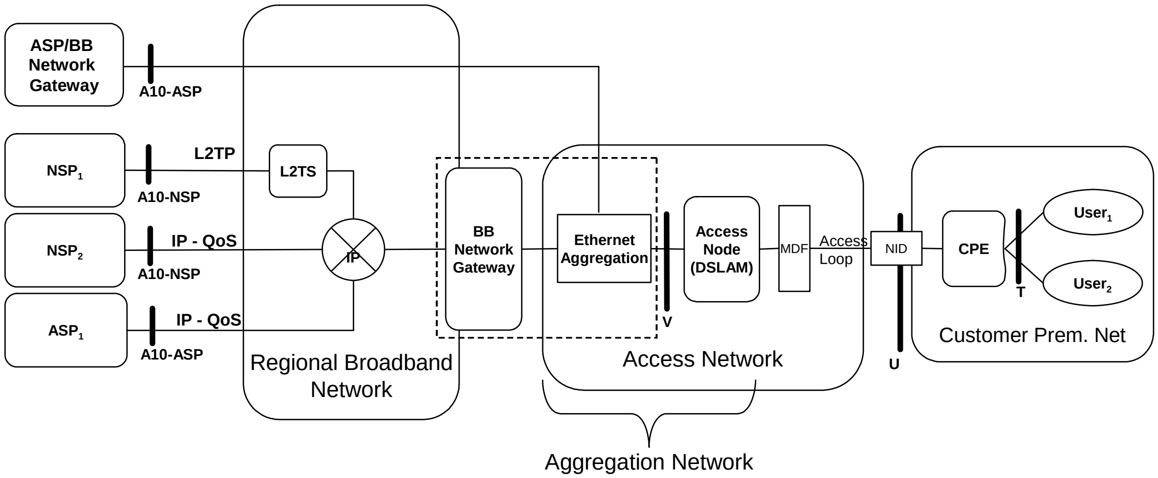


Figure 3 TR-101 Ethernet-based High Level Reference Architecture Example

4.1.2 QoE-based Engineering

The process of engineering a network for services includes:

- end user requirements analysis,
- definition of application layer QoE requirements
- translation from subjective QoE requirements to objective service performance requirements end-to-end at the network and application layers
- allocation of performance impairments to protocol layers, network segments or nodes

An example QoE engineering methodology is illustrated in Figure 4.

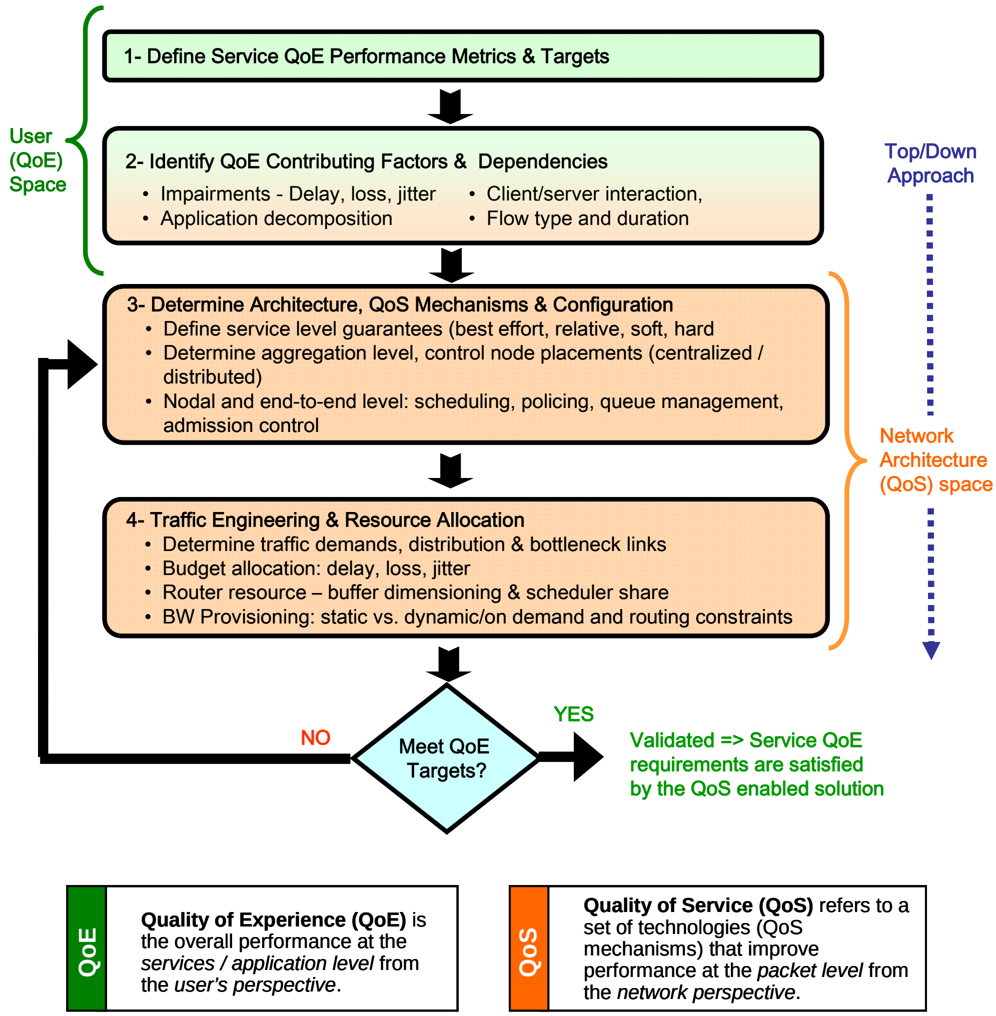


Figure 4 QoE Engineering Process

.

In the QoE engineering methodology, the inter-relationship of QoE / QoS and traffic engineering is defined based on a top-down approach, starting at the end-user level. The top-down approach is an effective technique to deliver enhanced customer value while performing network engineering. This document focuses on the first two steps of the process in defining the QoE requirements and corresponding service performance requirements.

4.2 QoE Dimensions

To identify factors that underlie QoE for a new service, three aspects of service delivery should be considered:

(1) The process of setting up a session (logging on, dialing a call, establishing a VoD session, etc.),
(2) How the service operates once the session is established
(3) Session teardown (logging off, shutting down).

Note some services may not be session based and in some instances may not be session aware, in which case the quality of experience will be limited to the service operation aspects.

In each of these aspects of using a service or application there are multiple facets that contribute to the overall quality of experience including:

- Effort required by the user
- Responsiveness of the application / service (Control Plane & Data Plane)
- Fidelity of information / content conveyed (Data Plane)
- Security / trustworthiness
- Dependability / availability.

When examining a complete service environment in any of these planes, there are three layers to consider:

- Service Layer
  - The layer exposed to the user, where QoE is measured
- Application Layer
  - Where various application parameters are set, for example media resolution, codec type, bit rate, decoder loss concealment, application layer error correction mechanisms, etc.
- Transport Layer
  - Where various impairments (loss, delay, jitter) may occur and where QoS and error correction mechanisms may be employed

This Technical Report provides the following information for each service:

- Identify the variables contributing to satisfactory service QoE at the service layer for control plane, data plane, and dependability dimensions
- Provide recommendations for Application Layer variable values required for satisfactory QoE
- Provide recommendations for transport layer variable values to achieve target QoE
- Define measurement methods to verify QoE.

The following sections provide definitions and listing of applications to be profiled.

4.2.3 Video – Broadcast and On Demand to Television Terminals

There are many types of video service:

- Video conferencing - real time, interactive, bidirectional video, room-to-room
- Video telephony – real time, interactive, bidirectional video, device-to-device
- Broadcast content video sources (TVoDSL) to Living Room TV - such as those traditionally distributed over the air or via cable or satellite Broadcast content
  - real-time (near), non-interactive, unidirectional
- Specialty or premium video content such as channels traditionally distributed via cable or satellite by pay-per-view, subscription and/or in specialty subscription bundles
  - real-time (near), non-interactive, unidirectional
- Video on demand (VoD) – on-demand delivery, also near VoD, network Personal Video Recorder (nPVR), CPE-based Personal Video Recorder (PVR), etc.
  - Real-time, interactive (with VCR-like controls), unidirectional video
- Security Applications – Surveillance
- Internet streaming to PC desktop or wireless device (cell phone, PDA, personal video terminal)
- Broadcast contribution and TV production networks (professional video)

each of which will have unique QoE requirements. There are also a multitude of video frame rates / resolutions and interlaced (i) versus progressive (p) scan to consider including:

- North American 29.97 frames per second (fps) / 59.94 fields per second and European 25 fps / 50 fields per second for interlaced materials
- Standard Definition resolution (North America 480i, 480p / European 576i, 576p) and many variations of SD including DVD vs. broadcast, half, three quarter and full horizontal resolutions
- High Definition – 720p, 1080i, 1080p

Consumers will compare broadband video service quality to alternative delivery mechanisms including CATV, over the air digital (ATSC in North America, DVB-T in Europe), digital cable, digital satellite, and DVD sources.

This Technical Report provides QoE objectives for entertainment video:

- Broadcast, specialty or premium video content source services delivered over an broadband infrastructure either as multicast or unicast streams
- Video on demand (traditional movies, network PVR, time shifted on-demand broadcast content, PPV special events, etc.)
- SD resolution - SMPTE-125 480i (59.94), 480p, ITU-R BT.601 / BT.656 576i (50), 576p)
- HD resolution (720p – SMPTE 296M, 1080i – SMPTE 274M)

4.2.4 Voice

As with video applications there are many variants of ‘voice’ services including:

- Wired conversational voice (analog, TDM digital, IP digital, etc.)

- Wireless (cellular) conversational voice
- Wireless (cordless, single base station) conversational voice
- Voice messaging
- Interactive voice response (IVR) services,

Application and transport layer QoE performance objectives for digital conversational voice applications such as voice over IP (VoIP), and VoIP via analog terminal adaptors in the customer premise will be provided. Requirements for traditional telephony voice (Plain old telephone service – POTS) applications are well-understood and not included here.

4.2.5 Best-effort (BE) Internet Data

BE Data – This is the current way of delivering most service offerings over broadband infrastructure and includes web browsing, e-commerce, email, file transfer, streaming media, VPN, P2P, IM, etc.. Since the transport service is best-effort, and services could be provided from outside the transport service provider's control, no guarantees can be made. However, from an end user perspective, target QoE requirements can be established for these applications and services. This Technical Report has identified ITU guidelines in G.1010 to generally classify and provide performance guidelines. As these applications move from BE to guaranteed service levels, refinements to the targets can be made in future versions of this Technical Report.

4.2.6 Future application extensions

Others applications depending on resources and demand could be added to this or more likely future working texts including:

- Video conferencing
- Generic lower end video for other applications such as PC streaming, handheld devices, video mail, etc.
- Streaming video to PC for corporate / business use
- Audio conferencing including audio chat with gaming
- Secure data (VPN)
- FTP file transfer
- P2P
- Others to be determined by the DSL Forum Marketing group or others
- Etc.

5. Entertainment Video QoE Overview and Measurement Guidelines

The Entertainment Video application category comprises video on demand (VoD), broadcast video and premium video content (e.g. pay per view) services. These services include video streams that are of SD or HD resolution and traditionally targeted to television terminals, usually via a set-top box (STB). Entertainment video includes:

- Broadcast channels
- Specialty or premium channels
- On-demand content including movies, time shifted broadcasts, network PVR, live and recorded special events such as PPV, etc.

There are likely to be some subtle QoE differences between these types of content, including variations in both customer expectations and content owner requirements. However the QoE dimensions and measurement techniques will be the same. There will also be common baseline minimum overall QoE targets with specific variations noted as required.

5.1 Video QoE Dimensions

At the video Service Layer, QoE dimensions include:

- Control Plane:
  - Interactive responsiveness (channel change delay, VoD and PVR / nPVR control responsiveness)
- Data Plane:
  - Video 'picture' quality
    - Many potential impact points on video quality in an end-to-end system
    - Impairments: blocking, blurring, edge distortion, judder, visual noise, incorrect image data due to loss, etc.
  - Audio quality
    - Also interaction of audio and video on overall media quality QoE
    - Media synchronization
- Usability
  - Service UI (set-up, finding content - EPG, PVR, remote control…)
- Dependability: Reliability / Availability
  - Session blocking ratio
- Security / Privacy
  - for user, Telco, and content owners
  - security impacts on other dimensions (e.g.. encryption / decryption delay)
- Content
  - mainstream, high quality, popular content is a key TVoDSL success component, particularly for Video on Demand (VoD) services

A generic high level end-to-end video services delivery system is shown in Figure 5.

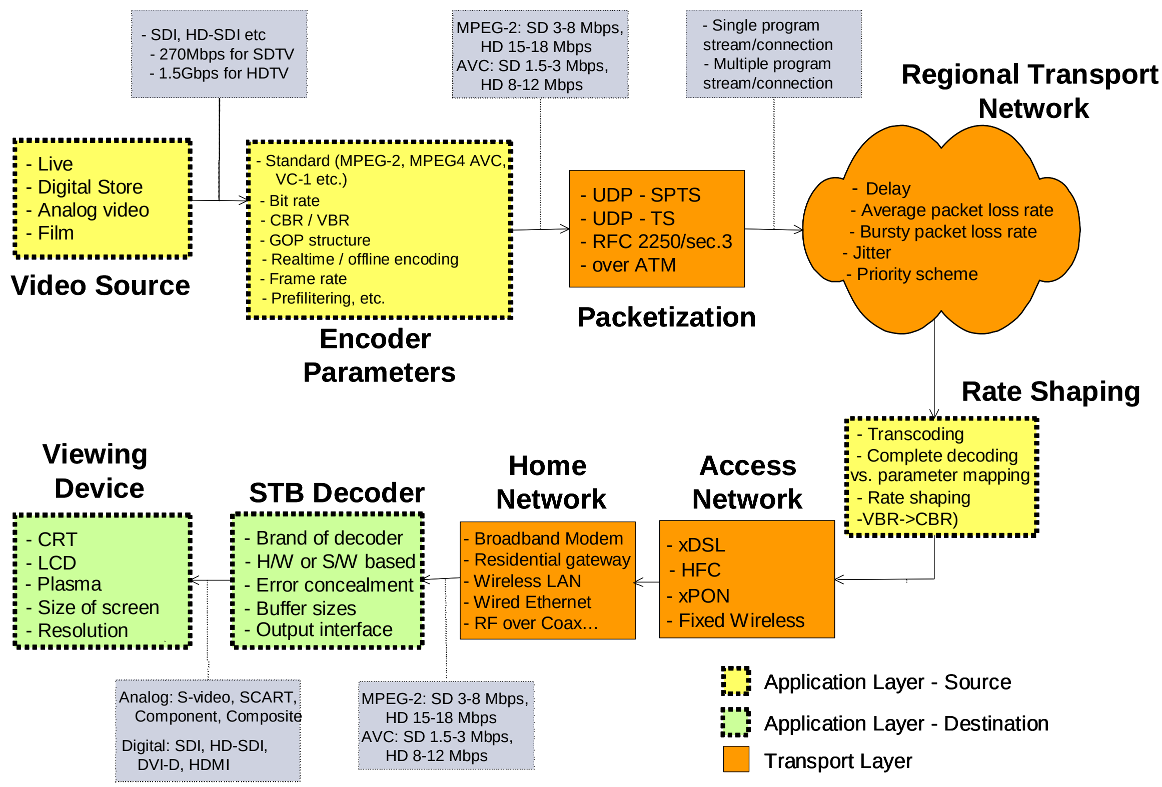


Figure 5 End-to-end Video Services Delivery System

As shown there are many contributing factors to end user QoE for a video service. At the video Application Layer factors include video acquisition and encoding (dotted line yellow boxes) and video decoding and display (dotted line green boxes), and any application layer error correction mechanisms (e.g. ARQ, FEC). Transport Layer elements include packetization and transport including regional, access and home networks (solid line orange boxes).

Video acquisition and encoding components are video source, video encoding, and rate shaping. The source of video can be a film, analog tape, digital storage (such as a VoD server), or live event (analog or digital). The quality of original materials greatly affects encoding efficiency and overall quality. Noise in the source materials wastes encoding bits and can affect quality. In addition, the source materials may be of varying resolutions and therefore varying quality to begin with. Video encoding is accomplished using video codecs suited for the particular transmission method and capacity. Depending on the type of application, several parameters of video encoding are defined, including bit rate, Group of Pictures (GOP) structure, constant or variable bit rate (CBR/VBR), and frame rate. Currently for broadcast applications, MPEG-2 is widely used. However, MPEG-4 AVC (also known as MPEG-4 Part 10, H.264, JVT) and/or SMPTE VC-1 (Windows Media 9) are expected to gain market share as their bit rates improve (reductions of two times are anticipated eventually) for comparable quality over MPEG-2. Video rate shaping (also known as digital turnaround or grooming) is required in the deployment scenario where the access network data rate is lower than the original source video coded bit rate or when the access network has links with differing bit rate capacity. Rate shaping sometimes includes a transcoding step where MPEG-2 is re-encoded in MPEG-4 AVC or VC-1 and can also convert a variable bit rate (VBR) stream to a constant bit rate (CBR) or capped VBR stream to facilitate network engineering and constrain bandwidth requirements. A content delivery service provider (e.g. Telco) may also receive the content streams pre-encoded by a third party aggregator for use directly in their network (e.g. no transcoding required) and may want to monitor the incoming quality to ensure it meets the requirements contractually agreed to.

Video packetization and transport components are shown in orange boxes (packetization, packet network, access network, home network). Packetization occurs at both the MPEG level and the network transport level.

At the MPEG level, video programs can be packaged individually as Single Program Transport Streams (SPTS) or in groups as a Multiple Program Transport Stream (MPTS), each with MPEG transport packets of 188 (originally chosen for ATM compatibility) or 204 (additional 16 bytes of error correction data) bytes. SPTS is used in Telco IPTV applications where only a single channel per TV is sent to the home due to access network bandwidth restrictions. . Note that an SPTS may include multiple audio streams e.g. for different formats and different languages etc. The aggregate bandwidth of these audio streams is not insignificant. In a bandwidth constrained IPTV application an operator may wish to send only the audio streams that are required. MPTS is used in digital cable and satellite applications where all content is broadcast to each home simultaneously.

At the network transport level, the MPEG SPTS or MPTS streams are then further packetized in the format required for transport. Typically in Telco IPTV deployments, IP transport is used with 7 MPEG packets per IP packet from the video head-end (as recommended in ProMPEG Forum Code of Practice #3[16]), and IP over ATM (AAL5) or Ethernet is used in the access network. Alternatively, MPEG-4 AVC / H.264 video streams may be packetized over RTP as proposed in RFC3984 - RTP Payload Format for H.264 Video[28].

In packet networks, issues of delay, jitter, and loss must also be dealt with. For broadcast TV applications, delay variations (jitter) are generally not problematic as there typically is a buffer in the set-top box (STB) at the customer premises. This de-jitter buffer adds delay in the range of 100 to 500 ms to remove time-of-arrival variation (packet jitter) from the data as it arrives. However, video quality degrades severely with packet loss, as well as with the type of MPEG information lost. Priority-marking algorithms (e.g. DiffServ) can be used to protect video quality from congestion losses as video traffic passes through the network.

The access network may consist of co-ax, copper, fiber, or fixed wireless network elements. A minimum of 12 to 24 Mbit/s is required to offer video services to the home based on 2 SD TV channels and 1 HD channel plus BE Internet and voice data. Currently, broadband technologies (bonded ADSL2+ and VDSL) seem to be the most economical means of deployment of real-time streamed video services which include a package supporting HD channels. For DSL access technologies, the data rate is a function of the copper loop length, with limitations resulting from crosstalk from neighboring copper pairs in a bundle, as well as from other noise sources (e.g., RF interference, disturbance due to lightning, etc).. A significant challenge is to ensure that the FEC inter-leaver depth is more than the duration of impulse noise to achieve satisfactory BER without impacting other services. Other significant challenges include resynchronization events for the DSL modems and protection switching events in the access infrastructure. Detailed operational error event profiles are network specific. ANSI/TIA 921[29] provides some guidance on the expected error performance for different categories of networks.

The home network is another potential source of video impairment and is less well controlled (from the service provider perspective) than the rest of the network. Often video distribution in the home will be done using a separate physical network to isolate other home traffic from the video stream. The existing co-ax cable used for analog video distribution in the home is typically targeted with Ethernet over co-ax and analog RF used.

Any packet loss specifications set to ensure video service QoE must be set from an end-to-end perspective – from the video head end to the set-top box output including losses due to late arriving packets and any loss protection mechanisms used.

Video decoding and display components are shown in blue boxes (decoder, viewing device). Video decoding is typically done by set-top box (STB) hardware, which also performs program stream de-multiplexing and clock synchronization. Buffering in the STB (designed to compensate for jitter in the packet arrival time) and error concealment algorithms employed at decoders are important contributors to the resulting video quality at this stage. However, increased buffer size can degrade interactive functions such as channel change time or VoD controls, depending on how these are implemented. Error concealment algorithms employed at decoders remain mostly proprietary.

The following CPE factors influence video quality but are beyond the scope of this document. The interface between the STB and television can also impact video quality. STB to TV interfaces ordered from best quality to worst (typically) include digital, component, composite and RF modulated on a channel (e.g. channel 2 or 3). Video display devices can be a significant factor in an end user's perception of video quality. Issues such as type of screen (CRT, LCD, plasma, etc), size of screen, and image resolution can all affect perceived video quality. For example, the pixelization effect, as well as other impairments, are generally considered tolerable, if noticeable at all, on a standard TV, but to be more pronounced and objectionable when viewed on a large screen/high resolution TV.

5.2 Video QoE Measurement

The video picture quality contributions to QoE can be measured in three ways:

1. Subjectively using a controlled viewing experiment and participants who grade the quality using rating scales such as Mean Opinion Score (MOS)

2. Objectively at the service layer – using electronic test equipment to measure various aspects of the overall quality of the video signal (e.g. PSNR)
3. Indirectly – using measurements of network impairments (loss, delay, jitter, duration of the defect) to estimate the impact on video quality, where there is an established relationship between QoE and QoS.

5.2.1 Subjective video quality measurement

The final arbitrator of video picture quality is the human viewer. The goal of any video delivery service is to please the customer with high quality images and service. Customers are becoming increasingly sophisticated judges of video quality with DVDs serving as the reference coupled with a migration to larger, higher resolution TVs. Subjective evaluation using human viewers to rate the video quality can provide the most accurate assessment of video quality from the perspective of a service provider's customers.

5.2.1.1 Important parameters for human perception of video quality

The human vision system is extremely complex and many properties are still not well understood. The matter is further complicated with video sequences where the system has both a spatial and a temporal dimension and interactions between the two. In general, viewers' assessment of picture quality depends on many factors including viewing distance, display size and resolution, brightness, contrast, sharpness, color saturation, naturalness and distortion. To complicate matters, there is a difference between objective fidelity / accuracy and perceived quality. For example, viewers generally prefer more vibrant colors even though they are not accurate or necessarily natural (de Ridder et al., 1995)[30].

Viewers are particularly sensitive to artifacts of digital compression and network-induced digital loss impairments that result in video distortion and unnaturalness. These artifacts generally look very different to degradation that occurs under analog transmission (fading, ghosting, sharpness, etc.) or natural conditions (distance, laze, light conditions, etc.) that the visual system has adapted to. Additional details of the human visual system and impact on video quality assessment can be found in Winkler (1999)[31].

5.2.1.2 Subjective test environment and methodology

Subjective evaluations are done either informally or using formal techniques. Informal evaluation of video quality is often done by service provider craftsperson on site and technical experts ("golden eyes") in the video system head end or during commissioning. These skilled experts often have years of experience in knowing what to look for. Unfortunately these "golden eyes" may not always be available, may not provide repeatable results, and may not reflect the preferences of the service provider's customer population.

Formal subjective evaluations use tightly controlled environments and many carefully qualified experiment participants who view various video clips and rate the quality. Generally television subjective video picture quality tests are performed following the guidelines established in ITU-R Recommendation 500 (ITU-R, 2002)[32] known as "Rec 500". Rec 500 provides detailed guidelines for standard viewing conditions, criteria for selection of subjects and video test sequences, assessment procedures and methods for analyzing the collected video quality scores.

There are a number of subjective video quality methods suggested in Rec 500 and selection of the method used requires careful consideration of the impairments being evaluated, sequence duration, desire to closely model a home viewing experience (ecological validity) and other factors. The output of the subjective tests is often an average of the quality ratings called a Mean Opinion Score (MOS). The Video Quality Expert Group (VQEG) provides additional details in the various test plans on their web site (VQEG, 2004)[15].

When done properly, formal subjective tests provide accurate and ecologically valid assessment of video picture quality. Unfortunately formal testing is time consuming and requires a specialized facility, making it expensive. In addition, it can be a complex task to design and run a subjective test that provides statistically interpretable and repeatable results. Even within Rec 500 there are many possible test methodologies and procedures to choose from, depending on the goals of the experiment.

The complexity, time and costs of subjective video quality evaluation has driven many researchers to attempt to create models of human video assessment in an attempt to replicate the scores given by subjects in an objective tool. These so called Perceptual Video Quality Measurement (PVQM) algorithms have undergone significant improvement in recent years but much remains to be done as outlined in the following sections.

5.2.2 Objective video quality measurement

As shown in Figure 5, and discussed in the previous sections, there are many parameters and components affecting video quality in IPTV systems. To assess video quality of the IPTV environment, it is important to efficiently conduct the video quality tests in timely fashion. Objective video quality measurement techniques, although not as accurate as subjective video quality measurement, offer a good compromise to conduct video quality assessment tests. Objective video quality measurement tests can be performed quickly to support fine tuning of network variables.

In recent years many advances have been made in objective quality monitoring techniques. While objective measurements with good correlation to subjective quality assessment are desirable in order to attain optimal quality of experience in the operation of broadband systems, it must be realized that objective measurements are not a direct replacement for subjective quality assessment. Objective measurements and subjective quality assessment are complementary rather than interchangeable. Where subjective assessment is appropriate for research related purposes and confirming objective measurements, objective measurements and/or indirect measurement methods are required for equipment specifications and day-to-day system performance measurement, improvement, and monitoring.

5.2.2.1 Objective video quality measurement techniques

Objective quality measurement techniques can be broadly classified into four categories:

- Techniques based on models of human video perception
- Techniques based on video signal parameters
- Techniques based on network impairment parameters
- Techniques based on the duration of the impairment in the video signal.

The traditional performance measurements of video transport and storage systems use fixed test signals and assume that the system under test is time-invariant. While these signals and the associated measurements are indispensable for the characterization of the electrical performance of conventional, time-invariant, analog video systems, the measurements often do not correlate well with video quality perceived by the end users of the video system. Following sections describe each of the methods, for digital video objective quality measurement, highlighting their strengths and weaknesses.

5.2.2.2 Measure Using Model of Human Video Perception

Objective quality measurement techniques falling into this category attempt to emulate the characteristics of the human vision system to obtain the video quality scores that have high correlation to the ratings actual viewers would provide. The human visual system modeling methods can be using one of the following approaches:

- 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.

5.2.2.3 Measure Video Signal Parameters Directly

In this method video signal characteristics are used to compare transmitted and received video streams. One way is to do a frame by frame, pixel by pixel comparison of the two video streams and calculate a mean square error (MSE) between the two. The difference between two video sequences can also be expressed as a picture or peak signal to noise ratio (PSNR). PSNR is calculated by the log of the ratio of the peak signal squared to the MSE in a similar fashion to analog systems. Until the emergence of models using human perception, PSNR was a widely used method of comparing the video quality.

PSNR gives some measure of the difference between two video clips, however this method does not take into account the parameters important to human perception as described in the previous section. For example, applying pre-filtering prior to encoding can degrade the PSNR but improves perceptual video quality. PSNR scores are independent of viewer's sensitivity to and tolerance for signal degradation. In subjective tests performed by VQEG, PSNR was shown to have low correlation to viewer ratings of video quality. The limitations of PSNR should be considered when assessing video quality, particularly for newer encoder implementations and fine analysis.

5.2.2.4 Standards for Objective video quality measurement

A growing concern of video researchers and broadcasters alike is the assurance and maintenance of an acceptable service quality level for the distribution of video programming. In 1997, the ITU created the Video Quality Experts Group (VQEG) to address video quality issues, in particular the development and standardization of accurate objective methods for estimating subjective video image quality. Traditional analog objective measurement systems, while still necessary, are not adequate to measure the quality of digitally compressed video systems. With the shift in technology from analog to digital and from synchronous to packetized transport, the types of visual artifacts have changed. To properly assess these new artifacts, new objective methods need to be developed and standardized. Many of VQEG participants are also active in the ITU-T SG9 and SG12 and ITU-R SG6.

The goal of the VQEG is to evaluate and recommend objective video quality measurement techniques and feed the final recommendation to ITU and other bodies for standardization. Selection of an objective quality measurement technique for a particular application consists of four parts:

- Definition of the test conditions including test material, type of codecs, channel conditions, viewing environment, test results analysis criteria, etc.
- Proposals of objective quality measurement techniques including the executable code to perform quality measurements. The evaluation process is open to all creators of objective video test methods.
- To ensure fair testing of proposed methods an Independent Lab Group (ILG) defines and conducts the subjective testing and processing of data through the models. ILG members include Verizon/USA, CRC/Canada, Nortel/Canada, Intel/USA and FUB/Italy.
- Finally the subjective and objective testing results are compared, and correlation analysis undertaken to select best objective quality measurement technique(s).

Until recently, VQEG was active in evaluating the objective quality method for Full Reference TV (FR-TV). In January 2004, VQEG submitted the final FR-TV report to ITU including four recommended objective models for FR-TV quality assessment that demonstrated good correlation to subjective video quality ratings in the Phase II evaluations (VQEG Phase II, 2003). These double-ended measurement methods with full reference, for objective measurement of perceptual video quality, evaluate the performance of systems by making a comparison between the undistorted input, or reference video signal at the input of the system, and the degraded signal at the output of the system. Unfortunately the FR-TV tests did not include packet loss impairments and the applicability of the tools to evaluate digital packetized transports remains an open question.

Currently VQEG is focused on addressing objective quality measurements: RR-NR TV (reduced reference / no reference, standard definition), HDTV (FR/RR/NR), and FR/RR/NR Multimedia (despite its name, the multimedia tests are still limited to video media only (no audio) in the first phase but for lower bit rates, lower resolution and smaller screens than typically used in TV applications). These new tests will include packet network impairments and hopefully will identify some useful tools for the objective measurement of video quality in broadband deployments. Until these tests are completed, service providers must rely on other methods of video quality assessment. There are a number of other objective video quality products on the market. Unfortunately these have not been evaluated independently (by VQEG or an independent lab) but some may be candidates in upcoming evaluations.

5.2.3 Indirect - Measure Network Impairment Parameters

It is important to monitor the quality of compressed video transmitted over a packet network from the perspective of a network service provider. It is not always feasible to use full reference (FR) or even reduced reference (RR) methods in all locations of the network since the reference may not be available. In addition, the no reference (NR) method would be prohibitively expensive to deploy widely for ongoing performance monitoring when there are many independent video streams to monitor as each of the streams would require separate decoder.

The Network Impairment method uses packet network parameters such as packet arrival time, delay, jitter, loss, impairment duration, and sequence number to extrapolate the video quality. Once the other evaluation tools are used to set tolerance for network performance based on QoE goals, these network parameters can then be monitored to provide a relatively cost-effective indication of the contribution of network behavior to video quality.

As shown in Figure 6, network layer metrics (e.g. packet loss, packet delay) have a direct impact on service layer quality (e.g. PSNR) which may be nonlinear. Application or network layer techniques such as packet loss protection (e.g. FEC, ARQ) also impact the objective service layer quality. There may be multiple network layer quality metrics (e.g. loss, delay, impairment duration) that impact a given service layer metric.

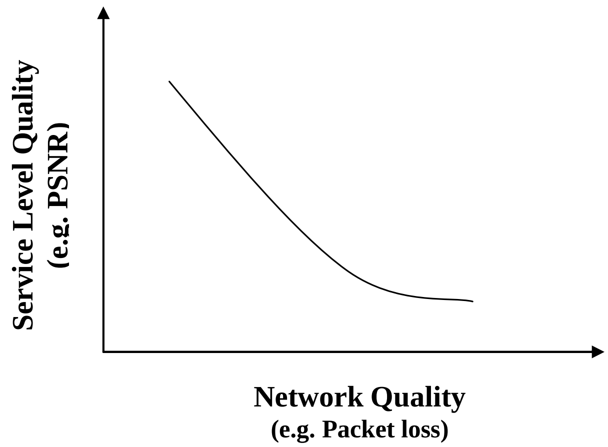


Figure 6 Relationship between Network Quality and Service Layer Quality

The relation between service layer quality and network layer quality metrics may be derived empirically if there is not a simple analytic approach available. Having identified the service/network layer quality relationship, it can be used in two ways:

- Given a network layer quality measurement, one could predict the expected service layer quality performance and QoE
- Given a target service layer quality, one could extrapolate the expected network layer quality performance requirements.

It should be noted that extrapolation of perceptual video quality based solely on network performance will never provide the entire video QoE picture. There are attempts at "snooping" the video traffic without fully decoding it, in combination with network parameters to assess video quality, but this is a very complex problem and it is expected to take some time before usable solutions appear in the market.

5.2.4 Video quality measurement recommendations

Given the costs and complexity of subjective video testing, objective methods are preferred for most instances. Objective video quality measurement techniques are in their infancy compared to voice quality evaluation tools (ex. E-model, ITU-T Recommendation G.107) but there are some promising advancements. Recently, evaluations of a few innovative objective video quality measurement techniques showing reasonably good correlation with subjective ratings have generated strong interest around the industry. Unfortunately, these algorithms employ a full-reference method to compute their estimates, which limits their applicability. As well, the evaluations did not include packet loss impairments, .and the algorithms have limited commercial availability.

Method	Strengths	Weaknesses
Human model – Full Reference, Reduced Reference, No Reference	Accurate, requires more research/standardization (for RR and NR)	Not feasible for in-service monitoring, may require transmitted video signal, and requires fully decoded video. Algorithms with high correlation to subjective testing have limited commercial availability.
PSNR	Simple	Low accuracy when compared with subjective quality assessment, requires decoded transmitted and received video signal, not feasible for in-service monitoring
Network parameters at the STB	Can be used for in-service monitoring, less complex to implement Through statistical inference and continuous collection of STB measurement reports can be used to assess overall network performance in the video distribution chain	Only evaluates network layer impacts, not compression layer effects or video source quality Source video quality should be known and measured as close to the Headend as possible to estimate the absolute quality delivered to the user..
MPEG or Video layer parameters at the STB	Can be used for in-service monitoring, less complex to implement Depending on the availability of certain parameters at the video player, these could be used to return measurement at the video application layer One example is the Frame rate which gives a rough estimation of the continuity of the service delivered. Another may be impairment duration. Note that at the STB other factors, such as buffer underflow , may prevent the video stream from being played back properly even if all packets were eventually delivered by the transport network.	Source video quality should be known and measured as close to the Headend as possible in order to estimate the absolute quality delivered to the user. Depends on player capability to offer these parameters.

Table 1 Objective Video Quality Measurement Techniques

Until more reliable and practical methods become available, combinations of several objective techniques are recommended including direct network layer performance measurement. In some cases, it may be useful to supplement objective quality monitoring data with subjective tests conducted by one of the independent subjective video quality labs. Table 1 lists the strengths and weaknesses of the major objective video quality measurement approaches.

At the time of publication, there are few practical quality measurement systems available that can be used to measure video quality in the lab or during field system qualification/commissioning with independently verified high correlation to subjective assessment. In the cases where there is access to the transmitted and received video signal, one of the commercial full reference FR tools that estimate subjective quality have been used in the broadcast industry and could continue to be employed in the Telco space with the caveat that correlation to subjective quality and analysis of packet loss impairments may not be fully accurate.

It's recommended that service providers adopt video quality measurement methods recommended by VQEG and standardized by ITU for example in ITU-T Recommendation J.144 Revised, "Objective perceptual video quality measurement techniques for digital cable television in the presence of a full reference", and ITU-R Recommendation BT.1683, "Objective perceptual video quality measurement techniques for standard definition digital broadcast television in the presence of a full reference". The currently standardized quality algorithms are full reference (FR)_and have limitations as previously noted. The higher utility reduced reference / no reference and HDTV methods are currently being evaluated by VQEG, and it's recommended that the algorithms with high correlation to subjective assessment and appearing in future ITU recommendations be adopted. Ideally methods to measure / estimate service quality should be deployed in network elements including the STB. Work is currently underway in DSL Forum (DSLHome™ Technical Working Group), ATIS IIF and ITU FG IPTV on defining the RG and STB-based quality measurement methods.

6. Entertainment Video Service QoE Objectives

This section outlines the recommended QoE objectives for video services in the control, data and dependability planes. Objectives are provided in terms of engineering measures at the application and network layers for standard definition and high definition video services (VoD and broadcast).

As indicated previously there are several layers (Service, Application and Transport) that may impact the user's quality of experience. In addition, each of these layers may have several dimensions or planes. Table 2 illustrates the Layers and Planes addressed by this document for triple-play services.

	Service	**Application**	**Transport**

Control Plane	X	X	X
Data Plane	X	X	X
Usability Plane			
Content Plane		n/a	n/a
Reliability Plane	X		
Security Plane			

Table 2 Planes and Layers Discussed

Usability plane (e.g. user interface, electronic program guide, etc.), Content plane (e.g. types and subject matter of content offered), and Security plane (user, operator and content protection) are out of scope. References are made to related industry activities in video quality including ITU-T J.241[13] and ATIS PRQC contributions on Enhanced IP-Based Video QoS Performance Objectives[33]

6.1 Video Service Layer Video QoE Guidelines

Service layer QoE metrics are typically measurements of user opinion of the service quality such as Mean Opinion Score (MOS) or objective estimates of viewer ratings. For data plane aspects, media fidelity is measured and QoE may also be expressed in terms of tolerable lower layer impairments such as Application layer bit rate and Transport layer packet loss. Service level QoE metrics may also be extrapolated from MOS to tolerable delay for control plane functions. The frame of analysis is the complete system end-to-end from the ingress of video content in the service head-end to the display on the viewing device in the customer premise. The performance reference points will be locally implemented (no transport layer) CPE-based applications such as playing a DVD. User experience in both video quality and control responsiveness of their DVD players will set the bar for any network delivered services such as VoD.

6.1.1 Video Service Layer Control plane QoE Recommendations

Guidelines will be provided for maximum tolerable latency for user control actions such as:

- Stream Control:
  - Channel change speed and scalability with load
    - Need to be able to handle very large peaks in multicast processing (channel changes – multicast leaves and joins) during commercial breaks or at start of major programs (sporting events, movies, prime time, etc.)
    - Needs to be competitive with other TV service offerings
  - VoD Control – remotely stored content
    - responsiveness of VoD pause, play, rewind, fast forward controls
    - should emulate local VCR-like speeds
- System start-up
  - may have large number of users turning on TV and STB at same time requiring network based STB initialization and authentication

- User Interface:
  - EPG navigation responsiveness
    - interaction should be perceived as instantaneous

In general the control plane aspects of any service can be aligned with Interactive and Responsive classifications as defined in ITU G.1010 and shown in Figure 7.

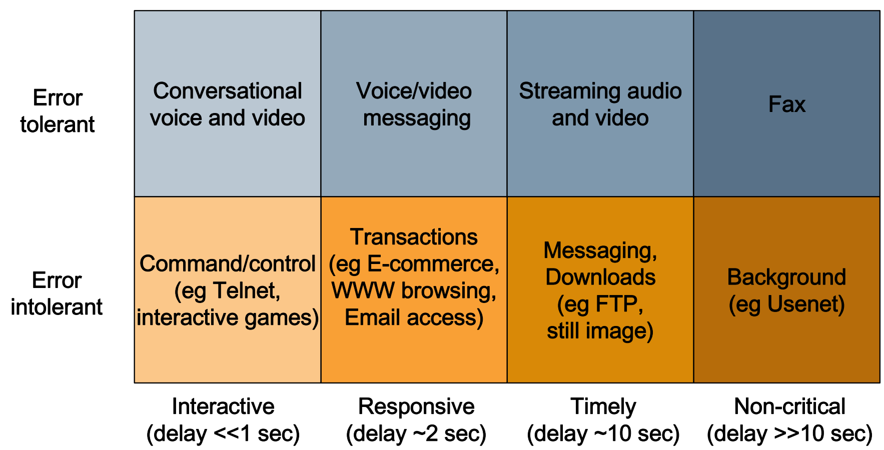


Figure 7 ITU G.1010 Application Delay Classifications

As shown in **Figure 7** for user interactions that should be perceived as taking place "instantly", such as UI interactions, the delay must be much less than 1 second – the Interactive class. A number of human factors studies have shown that for an action to be perceived as instantaneous, some feedback must be provided within 50 -200 ms [34],[35],[36]. Channel changing falls into the Responsive classification range with delays up to 2 seconds.

In addition control plane performance must be competitive with competing offerings from digital cable and digital satellite service providers. Typical digital cable channel change times can range from 1 – 2.5 seconds and digital satellite times in the range of 2 – 4 seconds. Table 3 lists the maximum recommended delays for control plane actions in order to achieve satisfactory user quality of experience. Of course service providers should strive to offer superior performance if possible, particularly for channel change where a target of 1 second is preferred. In addition, consistency of channel change response time is an important contributor to QoE.

User Action	**Examples**	**Category**	**Maximum Recommended**

			Delay
User interface actions	**EPG scrolling. VoD remote control button push to onscreen indication that command was received (ex. pause symbol displayed)**	**Interactive**	**200 ms**
Channel Change	**Time from remote control button push until stable channel is displayed on TV**	**Responsive**	**2 s**
System start-up time	**Time from STB power on to channel availability**	**Timely**	**10 seconds**

Table 3 Video Service Control Plane QoE Recommendations

6.1.2 Video Service Layer - Data plane

At the services layer, in the data plane, the media quality of video and audio over broadband must meet the service objectives of the operator. The operator may choose to offer services of different service quality depending upon a number of factors including the capital and operational cost implications of the implied network quality requirements. Ideally, media quality should be at least competitive with other service delivery mechanisms and preferably superior.

VoD and premium content may mandate superior levels of quality compared to broadcast because of increased user expectations aligned with additional cost to the user. Also HDTV services may require lower application and transport layer impairment levels than standard definition services to yield satisfactory service QoE. Section 5.2 Video QoE Measurement provides guidelines on measuring video QoE at the services layer. Eventually a standardized, acceptably accurate method of video quality measurement will become available but until such time, the data plane requirements will be limited to defining the behavior of the application and transport layers. The Application Layer (e.g., bit rate, resolution, etc) and Transport Layer (e.g., delay, loss, jitter) performance requirements are provided in subsequent sections.

6.1.3 Video Service Layer - Dependability plane

Dependability of a service will impact the overall user quality of experience. The dependability refers to the availability / reliability / survivability of the overall service which in turn is composed of network element (ex. router), network (ex. fiber), and service elements (ex. set-top box, video server, etc.) . Dependability requirements are normally quantitatively expressed in metrics such as downtime or percentage of service availability (e.g. 99.999% or so-called “five 9’s”).

There is some inconsistency in the industry in differentiating between dependability and quality, but typically the distinction is based on the severity of the impact to the user.

Reliability (availability) may be distinguished from general service quality by whether the service rendering is merely distorted (quality impact) or not intelligible or unavailable (reliability / availability impact). For example for a video service, if the picture quality of the channel stream is distorted due to packet loss but the picture content is still intelligible this would be a quality issue. However if the channel to be viewed is black (video stream not available) or the picture is so severely distorted as to be unintelligible (severe packet loss, decoder failure), then this would be deemed a failure and impact dependability metrics. A further distinction is sometimes made with respect to the temporal duration of the event with services degradation for durations greater than ten seconds being deemed an outage and included in dependability metrics. If the service degradation duration is below ten seconds, it is defined as a quality issue. Note that such quality issues include packet loss-based picture quality impairments, service delay, application failures, ineffective control attempts, and unavailability events.

Metrics for service dependability must include all factors from originating broadcast signal, satellite downlinks from content owners, right through to STB mean time between failures (MTBF). Non-availability due to blocking (with a corresponding service blocking probability), network bandwidth limitations, and other factors should be considered in addition to reliability factors such as equipment failures, commercial power outages and cable cuts. Total availability / reliability can be distributed across lower layers (Application and Transport) elements as deemed appropriate by service providers and will not be specified in this document. Note that while much of the service provider's network infrastructure and application specific video service equipment may be expected to be deployed in redundant configurations, the DSL Line, and home network equipment including the set top box are expected to be single devices.

As a baseline, any broadband supplied video service should have competitive and preferably superior reliability metrics to similar competing services such as satellite or HFC (hybrid fiber co-ax) network delivery mechanisms. The traditional Bellcore specification for telephony physical plant availability of 99.99% or 53 minutes of downtime per year has also been adopted as a target for cable operators with HFC plants as shown in Table 4.

Application	Source	Availability (Typical annual)	Notes
PSTN	Bellcore	99.99%	Includes fiber, host digital terminal and

	TA-NWT-000418 and TA-NWT-000909 [37]		optical network unit. Does not include unavailability due to electric utility power failures. Access plant only, does not include switch, drop, or in-home failures.
PSTN	Telcordia / Bellcore[38]	99.94%	Complete end-to-end voice service availability target.
MSO-HFC	CableLabs® Enhanced Services Deployment Guide[39]	99.99%	Studies found that HFC plant itself was a critical contributor to availability but a well engineered HFC network could meet and exceed the Bellcore specification
MSO-HFC	Stuart Lipoff[40]	99.992%	Performed independent analysis of HFC plant and availability factors and determined downtime of 41 minutes per year was possible.
MSO-VoIP on HFC	CableLabs® [41] VoIP Availability and Reliability Model for the PacketCable Architecture	99.94%	Demonstrated how a VoIP service could meet Telcordia circuit switched performance.

Table 4 Service Availability Targets

In addition, VoIP services end-to-end availability metrics have traditionally targeted 99.94% annual availability. Until reliable data on video service availability metrics from competing mechanisms is available the voice service guidelines of 99.94% availability with similar constraints as outlined in the VoIP service factors could also be adopted for the video service. It could be argued that the video service does have an emergency criticality like a voice service, with the Emergency Alert System (EAS) used in some countries to provide community based alerts via text overlays on a TV channel. It has also been demonstrated by video service vendors that consumers are very concerned about availability of their video services and the services are often “always on” during prime time viewing hours making outages immediately obvious.

6.2 Video Application Layer QoE Guidelines

At the application layer, parameters of video content such as resolution, frame rate, encoder and decoder settings, transcoding, bit rate, etc. are defined. In addition, corresponding audio track recommendations are made. At the receiving end, application layer parameters such as loss concealment are discussed. Frames of reference include end-to-end, encoder and decoder, and display device. Application layer service performance metrics are necessarily service specific. The following sections provide a brief overview of the types of service level parameters that may be of interest because of their impact on the system architecture or user visibility.

6.2.1 Application Layer - Control plane

Application layer control plane guidelines pertain mostly to responsiveness of user controls such as channel change and VoD control functions, but also to system start-up and electronic program guide or other UI responsiveness. The application layer impacts on the control plane functions are listed below. Note: a network delay component for most application layer functions is discussed in §6.3.1 Transport Layer - Control plane).

Channel change speed and scalability with load

- Set-top box (STB) command processing - time interval between the remote control action (button push) and the transmission of the leave / join message to the network
- STB layer delay - time needed by the STB IP stack to process incoming packets and deliver the content to the MPEG decoder engine. This may also include conditional access / decryption processing
- STB jitter buffer delay - time until the STB jitter buffer reaches the fullness set point prior to the forwarding of the video signal to the decoder function
- MPEG decoder delay - time required for the decoding process including the system buffer delay

VoD Control

- Set-top box (STB) command processing - time interval between the button press on the remote control and the transmission of the RTSP control message to the VoD servers and to update the on-screen GUI to indicate which button was pressed (pause, stop, rewind, fast-forward)
- VoD Server delay in processing the RTSP commands and generating required media stream (ex. fast forward or rewind)
- STB layer delay - time needed by the STB IP stack to process incoming packets and deliver the content to the MPEG decoder engine. This may also include conditional access / decryption processing
- STB jitter buffer delay - time until the STB jitter buffer reaches the fullness set point prior to the forwarding of the video signal to the decoder function
- MPEG decoder delay - time required for the decoding process

System start-up

- STB boot time
- Middleware server initialization and authentication
- Possibly firmware update time

EPG user interface navigation responsiveness

- Set-top box (STB) command processing - time interval between the remote control action (button push) and GUI update
- May include middleware server processing time for some functions

For control plane functions, it is left to the service provider to determine how best to partition the Application Layer and Transport Layer contributions to overall delay in order to meet the minimum recommended Service Layer QoE targets listed in Table 3 Video Service Control Plane QoE Recommendations.

6.2.2 Application Layer — Data plane

The main component of application layer QoE in the data plane is that of digitization and compression of video and audio source materials and the various settings and parameters selected.

Since video compression schemes such as MPEG are lossy and an identical copy of the original can not be recovered, there are potentially negative impacts on video picture quality and therefore on viewer QoE. The main factors influencing video QoE at the application layer due to compression are:

- Quality of source material
  - "garbage in = garbage out"
- The baseline quality (no network impairments) of the codec standard used
  - there are a range of video codecs available, but typically television applications will use one of the following: MPEG-2, MPEG-4 AVC (also known as MPEG-4 Part 10 or H.264) and SMPTE VC-1 (previously known as VC-9, the standardized version of Windows Media™ 9)
- Resolution
  - Some systems reduce the horizontal resolution to achieve the target bit rates for example in SD the resolution maybe reduced to 'Half' or "Three Quarters" which produces a less sharp picture than 'Full' resolution
- Bit rate
  - During periods of high complexity (entropy) compression may leave visible artifacts if the bit rate is not sufficient
- Application layer video encoding - Constant bit rate (CBR) vs. Variable Bit Rate (VBR) at the encoder output
  - Video encoding is naturally variable bit rate but to simplify network engineering for Telco delivery systems, the video encoders are set to provide a constant bit rate (as averaged over some specified time period on the order of seconds).
  - VBR streams such as those used in DVD encoding have constant quality since the bit rate is allowed to vary to accommodate varying complexity of the source material
  - CBR streams have variable quality since there may be times when the bit rate is insufficient to accommodate the video complexity but CBR steams enable more straightforward traffic engineering and system design
- Encoder quality and settings
  - Group of Pictures (GOP) structure
    - Shorter GOPs improve quality but reduce the improvement to bit rate from compression.

    - Longer GOPs improve maximum compression ratio, but increase channel change time and the amount of damage a lost packet will cause.
    - Dynamic GOPs can be used to better handle scene changes and other effects but are not always implemented on STBs. In addition, dynamic GOPs can impact the variability of zapping latency and may complicate mechanisms to increase zapping speed considerably.
  - Motion Vector Search Range
    - Wider searches provide improved quality but at increased complexity and encoder delay
    - Large search ranges are required for high motion content such as sports
  - Rate Control
    - Mode decisions greatly affect the bit rate
    - Proprietary schemes are commonly used to gain competitive advantage
- Preprocessing (such as noise reduction)
  - usually proprietary and non-standard but can improve bit rate / quality tradeoff
- Tandem encoding and rate shaping (e.g. digital turnaround)

Video Compression Artifact Examples

Figure 8 illustrates several kinds of compression artifacts that are largely due to insufficient bits allocated resulting in too coarse quantization of DCT coefficients or motion vectors and/or otherwise poor motion estimation. Addition details of compression artifacts may be found in Wolf (1990)[42].

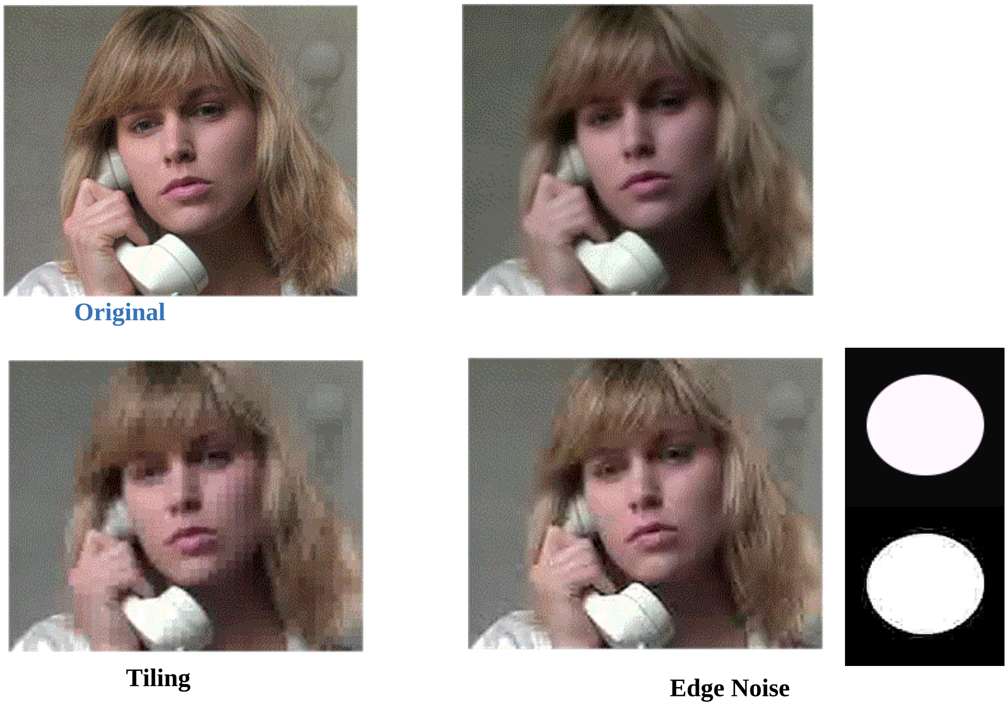


Figure 8: Compression Artifacts [ITS Video Quality Research (2003)[43]]

Similarly, there are multiple choices of audio codecs and similar parameter implications on the audio side. Most video service offerings (e.g. those using MPEG Transport Streams or similar) are capable of supporting more than one audio codec along with a single or sometimes multiple video encoding schemes depending on the headend equipment and set-top box. Commonly used audio formats for television applications include MPEG Layer II (also known as Musicam used in DVB systems, and MPEG-1, Audio Layer 2), Dolby Digital used in ATSC systems (formerly known as AC-3), NICAM 728 (European digital format for PAL), Advance Audio Coding – AAC (either MPEG-2 AAC or MPEG-4 AAC (ISO/IEC 14496-3, Subpart 4)), and sometimes other formats such as MP-3 (MPEG-1 Audio Layer 3) will be used, particularly for music content[44].

In addition to the separate audio and video application layer impairments, the synchronization between audio and video components must be maintained to ensure satisfactory QoE. There has been a great deal of research on A/V synchronization requirements in video conferencing and analog broadcast systems and specifications in such bodies as ITU-R[45, 46, 47]. Because audio that appears before video is very unnatural (sound takes longer to propagate than light so sound lagging visual is normal) some bodies specifying television specific A/V synchronization have recommended tighter tolerances than typically used for video conferencing applications [48].

Recommended minimum engineering objectives for application layer, data plane parameters are presented in the following sections for various video services. In general these parameters are guided by industry best practices (e.g. CableLabs® specifications, encoder vendor guidelines), performance of competitive systems (ex. cable, satellite benchmarks), telco deployment experiences (e.g. FastWeb), and the state encoding technologies (e.g. MPEG-2, MPEG-4 AVC, VC-1, etc. commercial offerings) at the time of publication of this document.

6.2.2.1 Standard definition (SD) TV: General minimum objectives

Table 5 lists the recommended minimum video application layer performance objectives at the MPEG level, prior to IP encapsulation for broadcast SD (480i / 576i). The audio stream bit rates are additional and specified separately below. Assumptions include:

Source material:
- 4:3 aspect ratio
- Source could enter the head end in analog or digital form

Maximum Viewable Resolution:
- Horizontal x Vertical: 720 pixels x 480 lines (North America) ITU-R BT.601-5 or 720 pixels x 576 lines (Europe)
- Lower resolutions (ex. ¾ Horizontal or ½ Horizontal – so called ½ D1) could be used to ensure encoding quality is maintained for complex materials

Frame rate:
- 29.97 fps (North America) or 25 fps (Europe)
- 23.97 / 24 fps may also be used for film-based materials (with 3:2 pulldown in North America for conversion to 30 fps)
- Two interlaced fields per frame

Video Codec standard	Minimum Bit Rate (video only)	Preprocessing Enabled
MPEG-2 - Main profile at Main level (MP@ML)	2.5 Mbps CBR	Yes (if available)
MPEG-4 AVC (Main profile at Level 3.0)	1.75 Mbps CBR	Yes (if available)
SMPTE VC-1	1.75 Mbps CBR	Yes (if available)

Table 5 Recommended Minimum Application Layer Performance for Standard Definition Broadcast Program Sources

Notes on SDTV Video Bit Rate

The bit rates achieved by a particular video encoder undergo continuous improvement over time particularly when first introduced. As is the case with MPEG-2 since its commercialization in the mid 1990s, improvements have typically followed McCann’s law that states encoder bit rate improves approximately 15% per year with the same quality[49].

In most cases the encoder improvements are done within the scope of the existing standards and therefore do not require upgrades to the decoders.

The MPEG-2 bit rates shown in Table 5 are nearing the end of the improvement cycle and although one may run at lower bit rates (particularly with proprietary preprocessing), the values indicated are the minimum required to provide adequate quality over a range of broadcast program material complexity. Note that many competing services (e.g., digital cable and satellite) use higher MPEG-2 bit rates and often VBR encoding. Where access link bandwidth permits, service providers are encouraged to use higher bit rates, particularly for broadcast materials with highly complex image content, such as sports.

MPEG-4 AVC and SMPTE VC-1 codecs are newer (broadcast systems became commercially available in 2005 for SD and 2006 for HD) and are similarly expected to improve over time, although perhaps not aggressively as suggested by McCann's law of 15% per year. The recommended minimum bit rate values shown in Table 5 represent the state of commercially available encoders at the time of publication. Table 5 assumes similar quality / bit-rate performance of MPEG-4 AVC and VC-1.

Table 6 lists the recommended minimum audio application layer performance guidelines for standard definition audio sources. Assumptions include:

Source material:
- NTSC (North America) or PAL (Europe / AsiaPac)
- Sources could include more than one stereo audio track to support multiple languages or multichannel audio for surround sound effects. Unless indicated in Table 6, recommended minimum bit rates are for one stereo pair only

Audio channels:
- Most broadcast content is now in stereo (left / right)
- Many broadcasters are also using Dolby 5.1 (up to 6 channels) for primetime series and special events, particularly concerts and sporting events

Audio Sample Rate:
- 48 kHz sample rate for Dolby digital as per ATSC
- 16 kHz to 44.1 kHz for MP3
- 32 kHz, 44.1 kHz or 48 kHz for DVB source audio as per ETSI TR 101 154

Audio Codec Standard	Number of Channels	Minimum Bit Rate (kbps)
MPEG Layer II	Mono or stereo	128 for stereo
Dolby Digital (AC-3)	5.1 if available, else left/right stereo pair	384 for 5.1 / 128 for stereo
AAC	Stereo	96 for stereo
MP3 (MPEG-1, Layer 3)	Stereo	128

Table 6 Recommended Minimum Audio Application Layer Performance for Standard Definition Sources

Notes on SDTV Audio Bit Rate

In general, audio codecs chosen should align with industry standards in the geography of deployment to ensure maximum compatibility with consumer receivers. There is a general trend to global support of Dolby Digital 5.1, particularly in North America (ex. ATSC) and this is also an option for DVB-based systems. Bit rates should be aligned with original source material quality and transcoding between formats should be avoided if possible. An MP3 target is provided to support music services.

Table 7 lists the recommended audio-video synchronization requirements based on guidelines provide by the ATSC for SD program materials[48]. Although these guideline were based on North American digital television they should apply equally well to formats in other geographies. Note the asymmetry in the requirement is due to the unnaturalness of audio leading video since light travels faster than sound.

Audio – Video Synchronization	Audio Lead Video	Audio Lag Video
	15 ms maximum	45 ms maximum

Table 7 SD Audio – Video Synchronization Requirements

Inconsistent loudness levels between channels can negatively impact QoE. It's recommended that equipment be used in the service provider head-end to ensure similar loudness levels across the range of channels provided to the user. Another audio quality issue beyond the scope of this document is the dynamic range compression for RF links between the STB and TV.

6.2.2.2 Standard definition (SD) TV: VoD and Premium Content Objectives

Video on demand (VoD) and other premium content such as pay per view in standard definition format will have similar application layer performance factors as regular broadcast materials. However, subscriber expectation may be higher because of additional fees paid to access the content and comparison to alternative delivery options. In the case of VoD, users may compare to VoD materials delivered over digital cable systems or even DVDs.

In North America, VoD application layer parameters are defined by Cable Labs®[19] . Since a great deal of existing VoD content is aligned with the parameters used by cable providers and consumers will compare the quality levels, it's recommended that telco-based video service providers adopt these as the minimum guidelines. The current guidelines are limited to MPEG-2 encoding. Recommendations for MPEG-4 AVC or VC-1 encoded VoD materials assume a 1.5x improvement in bit rate, aligned with the state of commercial deployments of these encoders. Table 8 lists the recommended video encoding bit rates for standard definition, VoD and other premium content and underlying assumptions below.

Source material:

- NTSC (North America) or PAL/SECAM (Europe / AsiaPac)
- 4:3 aspect ratio
- Encoding could be done offline using multipass systems for stored content such as VoD assets

Minimum Viewable Resolution:

- Horizontal x Vertical: 1/2 D1 352 pixels x 480 lines (North America) ITU-R BT.601-5 or 352 pixels x 576 lines (Europe) is permitted to ensure encoding quality is maintained for complex materials
- However, it's recommended that ¾ D1 resolution (528x480 / 528x576) be used where possible to align with the maximum specified for cable systems
- Telco service providers could run VoD assets at full D1 resolutions but would likely not be able to re-use assets pre-encoded for cable deployments

Frame rate:

- 29.97 fps (North America) or 25 fps (Europe)
- 23.97 fps may also be used for film-based materials (with 3:2 pulldown in North America for conversion to 30 fps)
- Two interlaced fields per frame

Video Codec standard	Minimum Bit Rate (video only)	Preprocessing Enabled
MPEG-2 - Main profile at Main level (MP@ML)	3.18 Mbps CBR	Yes (if available)
MPEG-4 AVC (Main profile at Level 3)	2.1 Mbps CBR	Yes (if available)
SMPTE VC-1	2.1 Mbps CBR	Yes (if available)

Table 8 Recommended Minimum Application Layer Performance for Standard Definition VoD and Premium Program Sources

Notes on SD Video Bit Rate for VoD and Premium Content

- Bit rates for MPEG-2 is as per Cable Labs® maximum for VoD content and aligns with the majority of VoD assets available[19]
- AVC and VC-1 bit rates are extrapolated from MPEG-2 using a 1.5x factor
- These guidelines are recommended minimums. Telco service providers are encouraged to use higher resolutions and bit rates where possible / practical for better quality
- Total video plus audio bit rate for most commonly available MPEG-2 encoded VoD assets is 3.75 Mbps
- The QoE of a VoD service may also be impacted by the quality of the implementation of trick mode features such as fast forward and rewind. The fast forward and rewind modes should be as smooth as possible and include intelligible audio during trick modes if possible.

Table 9 lists the recommended audio codec and bit rates for VoD and premium content. The bit rates assume a sampling rate of 48 kHz.

Audio Codec Standard	Number of Channels	Minimum Bit Rate (kbps)
Dolby Digital (AC-3)	5.1 if available, else left/right stereo pair	384 for 5.1 / 192 for stereo

Table 9 Recommended Minimum Audio Application Layer Performance for VoD and Premium Standard Definition Materials

6.2.2.3 High definition (HD) TV: Objectives

Table 10 lists the recommended minimum video application layer performance objectives for broadcast HD (720p / 1080i). Assumptions include:

Source material:

- ATSC (North America) or DVB (Europe) or TBD (AsiaPac)
- 16:9 aspect ratio
- Source enters the head end in digital form

Resolution and Frame rate:

- 720p60 (ex. SMPTE 296M) or 720p50 (DVB)
  - Horizontal x Vertical: 1280 pixels x 720 lines
  - 50, 59.94, 60 progressive scan frames per second
- 1080i60 (ex. SMPTE 274M) or 1080i50 (DVB)
  - Horizontal x Vertical: 1920 pixels x 1080 lines
  - 29.97 (59.94i), 30 (60i) interlaced frames per second, two fields per frame

Video Codec standard	Minimum Bit Rate (video only)	Preprocessing Enabled
MPEG-2 - Main profile at High level (MP@HL)	15 Mbps CBR	Yes (if available)
MPEG-4 AVC (Main profile at Level 4)	10 Mbps CBR	Yes (if available)
SMPTE VC-1	10 Mbps CBR	Yes (if available)

Table 10 Recommended Minimum Application Layer Performance for High Definition (HD) Broadcast Program Sources

Notes on HDTV Video Bit Rate

Bit rates for video encoding and corresponding decoders undergo continuous improvement over time particularly when first introduced. As with SD, improvements have typically followed McCann's law[50].

The MPEG-2 bit rates shown in Table 10 are nearing the end of the improvement cycle and although one may run at lower bit rates (particularly with proprietary preprocessing), the values indicated are the minimum required to provide adequate quality over a range of broadcast program material complexity. It should be noted that many competing services (ex. digital cable and satellite) use higher MPEG-2 bit rates and often VBR encoding. If access link bandwidth is available, service providers are strongly encouraged to use higher bit rates and/or VBR encoding for HD, particularly for complex broadcast materials such as sports.

MPEG-4 AVC and SMPTE VC-1 codecs are newer (broadcast systems commercially available in 2005) and are expected to improve significantly over time. The recommended minimum bit rate values shown in Table 10 represent the state of commercially available encoders at the time of publication but lower bit rates with satisfactory quality are expected as encoder technology improves. MPEG-4 AVC Main Profile is listed in Table 10, but as High Profile encoders and compatible STBs become available, service providers may choose to take advantage of superior features available for HD encoding in the High Profile. Table 10 also assumes similar quality / bit rate performance of MPEG-4 AVC and VC-1.

Table 11 lists the recommended minimum audio application layer performance guidelines for high definition audio sources, guided by industry best practices, performance of competitive systems (ex. cable, satellite), telco deployment experiences, and the state encoding technologies at the time of publication of this document. Assumptions include:

Source material:

- ATSC (North America) or DVB (Europe) or TBD (AsiaPac)
- Sources could include more than one audio track to support multiple languages
- For HD materials multichannel audio for surround sound effects should be provided where possible

Audio channels:

- Many broadcasters are also using Dolby 5.1 for primetime series and special events, particularly concerts and sporting events

Audio Sample Rate:

- 48 kHz sample rate for Dolby digital as per ATSC
- 16 kHz to 44.1 kHz for MP3
- 32 kHz, 44.1 kHz or 48 kHz for DVB source audio as per ETSI TR 101 154

Audio Codec Standard	Number of Channels	Minimum Bit Rate (kbps)
MPEG Layer II	Mono or stereo	128 for stereo
Dolby Digital (AC-3)	5.1 if available, else	384 for 5.1 /

	left/right stereo pair	128 for stereo
AAC	Stereo	96 for stereo
MP3 (MPEG-1, Layer 3)	Stereo	128

Table 11 Recommended Minimum Audio Application Layer Performance for High Definition Sources

Notes on HDTV Audio Bit Rate

In general, audio codecs chosen should align with industry standards in the geography of deployment to ensure maximum compatibility with consumer receivers. There is a general trend to global support of Dolby Digital 5.1, particularly in North America (e.g., ATSC) and this is also an option for DVB-based systems. Bit rates should be aligned with original source material quality and transcoding between formats should be avoided if possible. An MP3 target is included to support music services.

A/V synchronization requirements for HD materials is currently under study the ATSC and other bodies, Until additional data is available, the guidelines presented in Table 7 for SD materials should be followed for HD materials as well.

6.3 Transport Layer Video QoE Guidelines

Transport layer requirements are typically expressed using network performance metrics with appropriate targets and limits to meet the desired Service layer QoE. Parameters include bandwidth, transport stream packetization, network loss and/or error, latency, and jitter. The frame of reference is the complete transport system from head-end to STB including any application layer loss protection mechanisms (e.g. FEC, ARQ), national / regional, access and home networks and associated network elements. The transport layer performance guidelines provided in the following sections are with respect to the viewer experience and are measured after any application layer protection mechanisms employed to overcome the network impairments. Figure 9 illustrates the end-to-end frame of reference for the transport layer performance metrics (e.g. packet loss).

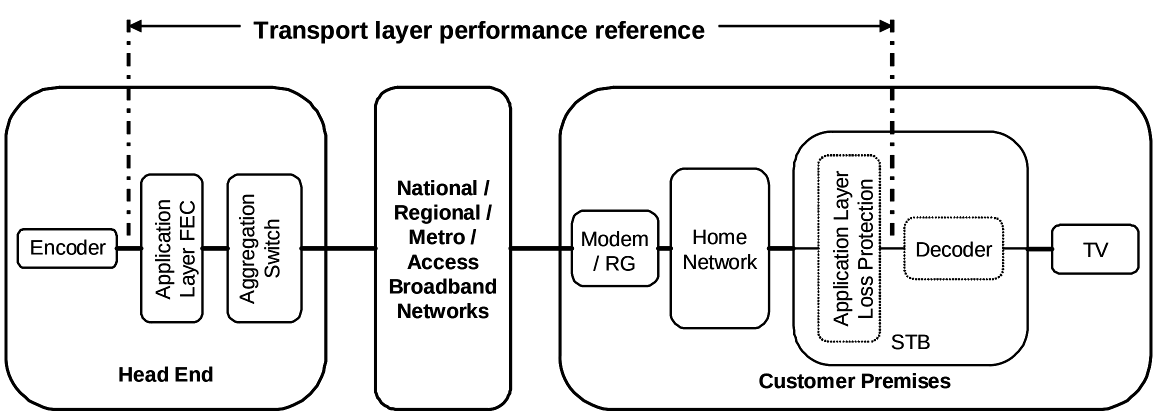


Figure 9 Transport Layer End-to-end Reference

Transport layer quality requirements can be approached from two different perspectives :

(i) the typical performance achievable from a managed IP/DSL network (e.g. ANSI/TIA 921 Profile A or B networks) and

(ii) the impacts of network defects like loss and delay on the service quality.

G.1010 also provides an approach to (i). The objectives set out in the tables in the following sections provide an approach to (ii). It is up to the application design / operator to resolve any differences e.g. through loss control techniques, design constraints on delays etc.

6.3.1 Transport Layer - Control plane

One of the important aspects of a good user experience for broadcast TV services includes fast channel change (also known as zapping). There are several factors that determine the zapping delay, i.e. the time between pushing the remote control button and the first video frame being rendered on the viewing device.

The figure below gives an example break-up of the channel zapping delay. The example is based on the following assumptions:

- The Routing Gateway and the Access Node both employ IGMP Immediate Leave (cf. TR-101)
- The DSL line is configured to use interleaving mode in downstream
- MPEG-2 encoding at 25 fps with 15/3 Group Of Pictures (GOP) structure

To ensure interactivity and satisfactory QoE, the channel zapping delay needs to be below 2 seconds as indicated in Table 3. The network layer impacts on zap time include join / leave time on the wire, time for IGMP processing at the control plane and time for the new stream to reach the STB. The application layer zap time factors were presented in Section 6.2.1 above. The example in Figure 10 shows both the network and application layer factors and that the application layer delay dominates the overall zap time. The zapping delay is primarily determined by the time require to have an I-frame available at the STB to start decoding the new channel, which in turn depends on the Group Of Pictures (GOP) length (distance between key or I-frames). This is because the I-frame is the only type of frame that does not need a reference to previous frames in order to be reconstructed correctly. If a zap occurs immediately after an I-frame, one has to wait a full GoP (which defines the time between two I-frames) + the time to transmit the following I-frame, before the channel can be played out error-free.

The delay can typically be 600 ms (e.g. GOP=15, 25fps) in case the previous I-frame has just been missed and an entire Group Of Pictures (GOP) needs to be sent. Therefore, a trade-off needs to be made between the GOP size and the encoding efficiency: the longer the GOP size, the more efficient the encoding (lower bit rate), but the larger the time between two I-frames (longer the channel change).

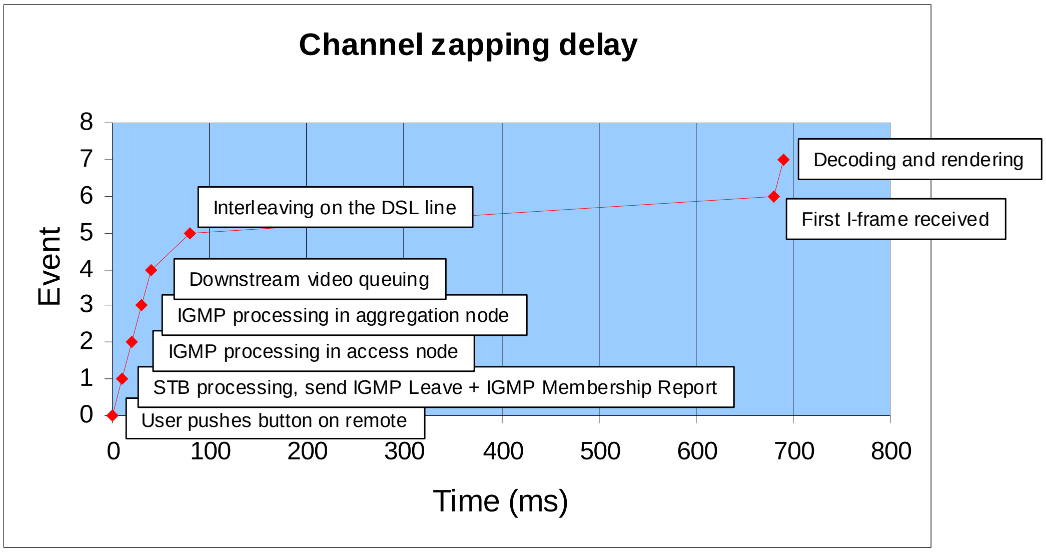


Figure 10: Channel Zapping Delay

Other factors that play a significant part in the overall channel zapping delay are the interleaving process on the DSL line (16-20 ms) that ensures protection against a certain level of impulsive noise, and the dejittering buffer present on the STB (10's to 100's ms).

Note that techniques exist to help to mitigate the channel zapping delay at the application layer. For instance, upon zapping to another channel, the STB could first show a dialogue box with the program name, time, channel, etc. This ensures that end users will never view a black screen while waiting.

6.3.2 Transport Layer - Data plane

Key criteria for the data plane in the transport layer include loss, latency and jitter. In general, reasonable end-to-end delay and jitter values are not problematic due to STB dejitter buffers, provided the dejitter buffer size is provisioned to match network and video element performance. Video streams however are highly sensitive to information loss and the QoE impact is in turn correlated to a number of variables including:

- Highly dependent on type of data lost
    - System information and header losses produce different impairments
    - Lost data from I and P frames produce different impairments than B frame packet losses due to temporal error propagation
- Dependent on codec used
- Dependent on transport stream packetization used
- Loss distance and loss profile
- With high encoding bit rates, the stream is more vulnerable to packet loss impairments
    - For the same packet loss ratio, impairments due to loss on a higher rate video stream occur more frequently (i.e., there are more visible errors per

unit time) simply because there are more packets per second transmitted and each one has the same probability to be affected.

- Decoder concealment algorithms can mitigate perceptual impact of some losses.

An error or sequence of errors in a video bit stream can cause effects ranging from no noticeable audio or video impact to the user to complete loss of the video or audio signal depending on what was lost and the robustness of the implementation. Figure 11 shows an example of the impact of a single lost IP packet (containing seven MPEG-2 packets) on a video frame if the lost information is from a B or an I frame. As indicated, since the I frame is a key frame used in the compression of subsequent P and B frames, the I frame impairment propagates in time across 14 frames of video or almost a half second (assuming 33 ms per frame). If the lost packet impacted a B frame, the impairment impacted only that frame with a duration of 33 ms.. Note that no loss concealment algorithms were running at the decoder.

Single B-frame IP packet loss

Single I-frame IP packet loss

Figure 11: Impact of Single IP Packet Loss (B Frame and I Frame)

The following tables show IP packet transport loss and jitter requirements to achieve satisfactory service quality targets.

Network latency and jitter should be engineered to closely align with set top box jitter buffer provisioning (wait time and buffer size) and overall network design and therefore may vary from implementation to implementation. Typical set-top box de-jitter buffers can store 100-500 ms (of SDTV) video, so network jitter must be within these limits and delay variation beyond these limits will manifest itself as loss. Increasing buffering also negatively impacts channel change latency so ideally the de-jitter buffers should be set as small as possible. Objectives outlined for jitter are based on experiences of operators and STB buffering capabilities.

Packet loss objectives are stated in terms of loss period and loss distance as defined in RFC3357 One-way Loss Pattern Sample Metrics[20]. Essentially loss distance is a measure of the spacing between consecutive network packet loss or error events; a loss period is the duration of a loss or error event (e.g. and how many packets are lost in that duration). The loss rates in the tables below are objectives designed to ensure satisfactory end user service level quality assuming no or minimal loss concealment. If the network infrastructure performance is below the required levels, service providers may make use of network level techniques (e.g. interleaving and FEC) and application layer mechanisms (e.g. loss concealment, application layer FEC, Automatic Repeat Request (ARQ)) as outlined in Appendix II - Error Protection Mechanisms Overview to achieve the required performance levels. In addition, the use of these techniques may provide an improved quality of experience over competing service offers.

Ideally the maximum loss period would correspond to one IP packet since even a single lost packet can result in a very noticeable impairment as shown in Figure 11. However, to account for possible use in a xDSL environment, including loop impairment behavior and FEC techniques available at the xDSL physical layer (RS, interleaving), we consider a loss period greater than a single packet.

Often random bit errors or minor amounts of congestion cause an isolated loss event with a loss period of one packet. DSL errors are different; interleaved Reed Solomon FEC codes are typically used at the DSL physical layer, and when these are overwhelmed by powerful impulse noise it causes an uncorrectable error at the output of the DSL decoder with a loss period greater than a single packet. Uncorrectable DSL errors typically wipe out an entire block of length equal to the interleaver depth.

A common configuration for DSL is an interleaver depth (i.e. FEC block duration) of 8 or 16 milliseconds. The corresponding loss period will therefore be 8 or 16 ms. Depending on the video bit rate this will correspond to a different number of lost video IP packets.

The recommended loss period is specified as less than 16 ms, which provides a balance between interleaver depth protection from impulse noise induced xDSL errors, delay added to other applications and video service QoE requirements to reduce visible impairments to on average one per 60 minutes for SD resolution video streams. The loss period will result in different numbers of packets being lost, depending on bit rate of the video stream as shown in the following tables. This maximum loss period objective is provisionally set until further studies allow better tuning of the maximum loss period allowed, in a xDSL environment considering all DSL variants, protection mechanisms, and optimum settings.

For DSL access cases, DSL modem resynchronization events imply a packet loss outage duration on the order of 10-20 seconds. An IPTV system would not be expected to maintain normal service through such an event. Such events might be considered a service outage rather than a quality defect.

The video application should be able to operate normally in the presence of normal operational defects. One such normal operational consideration is the operation of protection switching mechanisms in the network. SONET/SDH protection switching mechanisms may result in a potential packet loss duration on the order of 50ms. For some other protection mechanisms (e.g. MPLS fast reroute, fast IGP convergence) the potential packet loss duration can be longer, on the order of 250ms. Service providers are encouraged to add mechanisms to minimize or eliminate the visible effect of such protection mechanisms as these events cascade to a large number of subscribers.

Considering some other protection mechanisms the potential packet loss duration can be longer. For example, a complete reconvergence of the IP (IGP) routing table would imply potential packet loss bursts on the order of 30sec. An IPTV system would not be expected to maintain normal service through such an event. Such events can be considered a service outage rather than an in service quality defect.

The guidelines in the following tables are derived from deployment experiences (e.g. FastWeb), objective studies (e.g. Figure 11; Green et al (2001)[51]), subjective studies (e.g. Appendix I – Subjective Experiments in Video Quality

) and existing standards (e.g. ATIS[33], Bellcore[52], ITU-T J.241[13], ITU-T Y.1541[12], DVB, etc.). In general the following principles are applied:

All impairments are specified as end-to-end objectives (from video origin to the video output of set-top box to the television including any loss correction mechanisms that may be applied at network or application layers).

- Loss Distance of error events should be limited to at most one per 60 minutes for SD materials and one per 4 hours for HD. Error event is defined as a loss or corruption of a group of a small number of IP packets each containing up to seven MPEG packets of 188 bytes in length.
- There should be sufficient noise margin in the xDSL link to combat line noise and enough FEC interleaver depth to combat impulse noise in order to achieve required BER to achieve the packet loss objectives, without undue degradation for other services.
- Set-top box decoders should employ error concealment techniques to minimize impact of loss or corrupted video packets.
- Appendix II - Error Protection Mechanisms Overview provides additional details on access BER, FEC and ARQ mechanisms.

The goal is to minimize visible artifacts to as few as possible using a combination of network performance requirements, loss recovery mechanisms (e.g. FEC, interleaver) and loss mitigation mechanisms (e.g. decoder loss concealment).

6.3.2.1 Standard Definition Video: Broadcast TV Transport Layer Performance Objectives

Assumptions for Table 12 below:

- MPEG-2 codec,

- MPEG-2 transport stream,
- seven 188-byte packets per IP packet
- no or minimal loss concealment (tolerable loss rates may be higher depending on degree and quality of STB loss concealment)
- encoder output to after any application layer protection mechanisms at the customer premises
- metrics are for the IP flows containing video streams only, IP streams for other applications may have different performance requirements

Transport stream bit rate (Mbps)	Latency	Jitter	Maximum duration of a single error	Corresponding Loss Period in IP packets	Loss Distance	Corresponding Average IP Video Stream Packet Loss Rate
3.0	<200 ms	<50 ms	<= 16 ms	6 IP packets	1 error event per hour	<= 5.85E-06
3.75	<200 ms	<50 ms	<= 16 ms	7 IP packets	1 error event per hour	<= 5.46E-06
5.0	<200 ms	<50 ms	<= 16 ms	9 IP packets	1 error event per hour	<= 5.26E-06

Table 12 Recommended Minimum Transport Layer Parameters for Satisfactory QoE for MPEG-2 encoded SDTV Services

Table 13 lists the QoE performance objectives for MPEG-4 AVC or VC-1 encoded standard definition video materials. Assumptions for Table 13:

- MPEG-4 AVC or VC-1 codec,
- MPEG-2 transport stream with seven 188-byte packets per IP packet
- no or minimal loss concealment (tolerable loss rates may be higher depending on degree and quality of STB loss concealment)
- metrics are end-to-end from head-end encoder output to after any application layer protection mechanisms at the customer premises
- metrics are for the IP flows containing video streams only, IP streams for other applications may have different performance requirements

Transport stream bit rate (Mbps)	Latency	Jitter	Maximum duration of a single error	Corresponding Loss Period in IP packets	Loss Distance	Corresponding Average IP Video Stream Packet Loss Rate
1.75	<200 ms	<50 ms	<= 16 ms	4 IP packets	1 error event per hour	<= 6.68E-06
2.0	<200 ms	<50 ms	<= 16 ms	5 IP packets	1 error event per hour	<= 7.31E-06
2.5	<200 ms	<50 ms	<= 16 ms	5 IP packets	1 error event per hour	<= 5.85E-06
3.0	<200 ms	<50 ms	<= 16 ms	6 IP packets	1 error event per hour	<= 5.85E-06

Table 13 Recommended Minimum Transport Layer Parameters for Satisfactory QoE for MPEG-4 AVC or VC-1 encoded SDTV Services

6.3.2.2 Standard Definition Video: VoD and Premium Content Transport Layer Performance Objectives

The requirements for network performance of broadcast SDTV applications listed above should be followed for VoD and premium content services also.

6.3.2.3 High Definition TV: Transport Layer Performance Objectives

It is commonly agreed upon that ideally HDTV services meet a criterion of one visible impairment event per 12 hours or better. In the remainder of this section, we propose a value of four hours as the minimum Loss Distance for HDTV services, assuming that not all errors will result in a visible impairment, because:

- loss of B-frame information is sometimes below threshold of noticeability
- error concealment will be used with HD decoders

Table 14 below shows the loss period and loss distance for MPEG-2 HDTV under the following assumptions:

- MPEG-2 codec
- MPEG-2 transport stream with seven 188-byte packets per IP packet
- STB has some level of loss concealment
- encoder output to after any application layer protection mechanisms at the customer premises
- metrics are for the IP flows containing video streams only, IP streams for other applications may have different performance requirements

Transport stream bit rate (Mbps)	**Latency**	**Jitter**	**Maximum duration of a single error**	**Corresponding Loss Period in IP packets**	**Loss Distance**	**Corresponding Average IP Video Stream Packet Loss Rate**
15.0	<200 ms	<50 ms	<= 16 ms	24 IP packets	1 error event per 4 hours	<= 1.17E-06
17	<200 ms	<50 ms	<= 16 ms	27 IP packets	1 error event per 4 hours	<= 1.16E-06
18.1	<200 ms	<50 ms	<= 16 ms	29 IP packets	1 error event per 4 hours	<= 1.17E-06

Table 14 Recommended Minimum Transport Layer Parameters for Satisfactory QoE for MPEG-2 encoded HDTV Services

Table 15 lists the QoE performance requirements for MPEG-4 AVC or VC-1 encoded high definition video materials.

Assumptions for Table 15 below:
- MPEG-4 AVC or VC-1 codec,
- MPEG-2 transport stream with seven 188-byte packets per IP packet
- STB has some level of loss concealment
- encoder output to after any application layer protection mechanisms at the customer premises
- metrics are for the IP flows containing video streams only, IP streams for other applications may have different performance requirements

Transport stream bit rate (Mbps)	**Latency**	**Jitter**	**Maximum duration of a single error**	**Corresponding Loss Period in IP packets**	**Loss Distance**	**Corresponding Average IP Video Stream Packet Loss Rate**
8	<200 ms	<50 ms	<= 16 ms	14 IP packets	1 error event per 4 hours	<= 1.28E-06
10	<200 ms	<50 ms	<= 16 ms	17 IP packets	1 error event per 4 hours	<= 1.24E-06
12	<200 ms	<50 ms	<= 16 ms	20 IP packets	1 error event per 4 hours	<= 1.22E-06

Table 15 Recommended Minimum Transport Layer Parameters for Satisfactory QoE for MPEG-4 AVC or VC-1 encoded HDTV Services

The PLR in the range of 10^{-6} recommended for video services may require special error control techniques to achieve the target. Appendix II - Error Protection Mechanisms Overview provides additional details on access network BER and FEC performance and mitigation options.

The network layer performance objectives are summarized in the figures below. Figure 12 shows packet loss ratios as a function of bit rate and time between uncorrected loss events for isolated packet loss events. Points from Table 12 and Table 13 are plotted as representative of SD video with a loss distance of one hour between packet loss events. Points from Table 14 and Table 15 are plotted as representative of HD video with a loss distance of 4 hours between packet loss events. The figure assumes that each IP packet carries 7 MPEG data packets, each 188 bytes long. The plots implicitly assume that error statistics are stationary and time invariant.

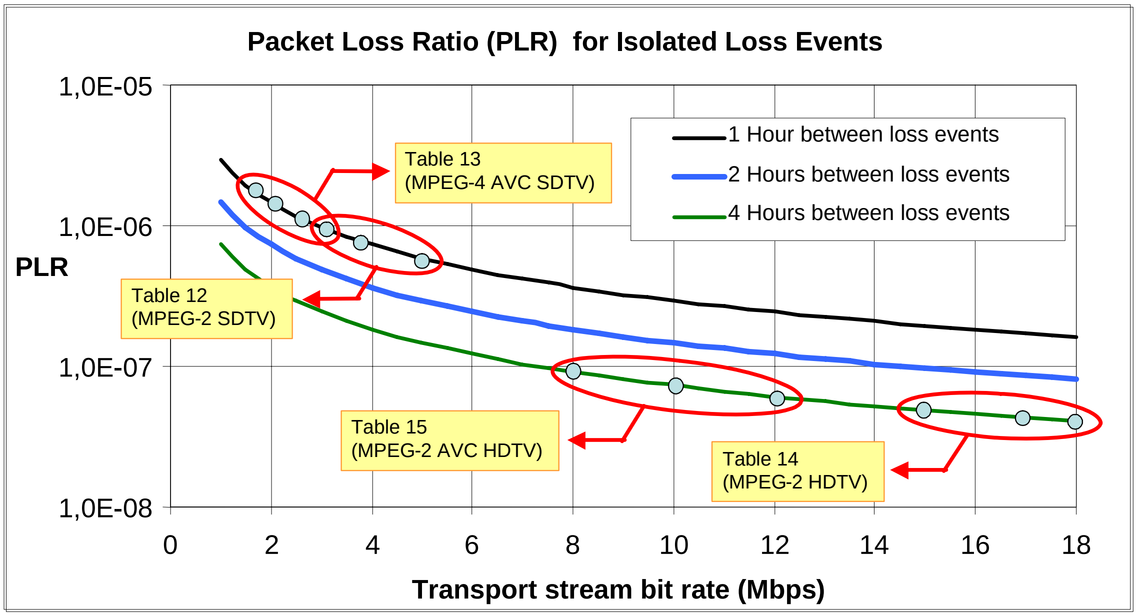


Figure 12: PLR required to meet average time between loss events of 1, 2 and 4 hours assuming isolated lost packets.

Figure 13 and Figure 14 show packet loss ratios as a function of bit rate and time between uncorrected loss events for typical DSL burst loss events of 8 ms and 16 ms, respectively. The "ripple effect" in the charts is the result of rounding to an integer number of lost/corrupted IP packets. For example, 8 ms of lost video data in an MPEG-2 transport stream at a bit rate of 3 Mbps:

Total MPEG packets per second = 3 Mbps / 8 bits per byte / 188 bytes per MPEG packet
= 1994.7 MPEG packets per second
Total IP packets per second = 1994.7 / 7 MPEG packets per IP packet
= 285 IP packets per second

A loss of 8 ms corresponds to = 285 IP packets per second * 0.008 seconds
= 2.28 IP packets lost.

Because an entire IP packet is lost if a part of a packet is lost, this is rounded to the next integer = 3 IP packets. And because the lost bytes are not necessarily aligned to IP packet boundaries, this would be further rounded to 4 IP packets.

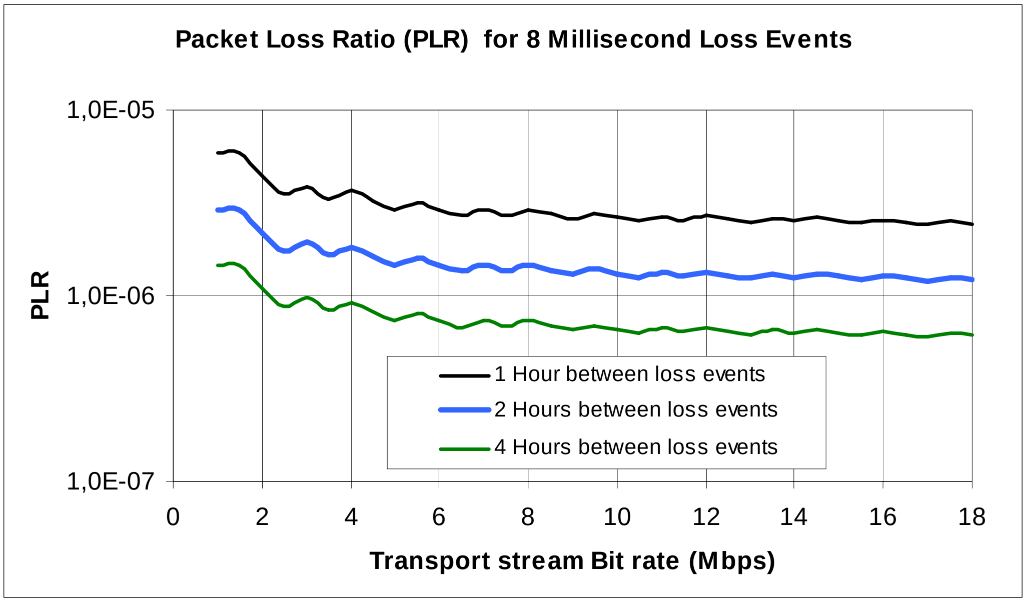


Figure 13: PLR required to meet average time between loss events of 1, 2, and 4 hours assuming each event is an uncorrectable DSL error that loses 8 milliseconds of contiguous data.

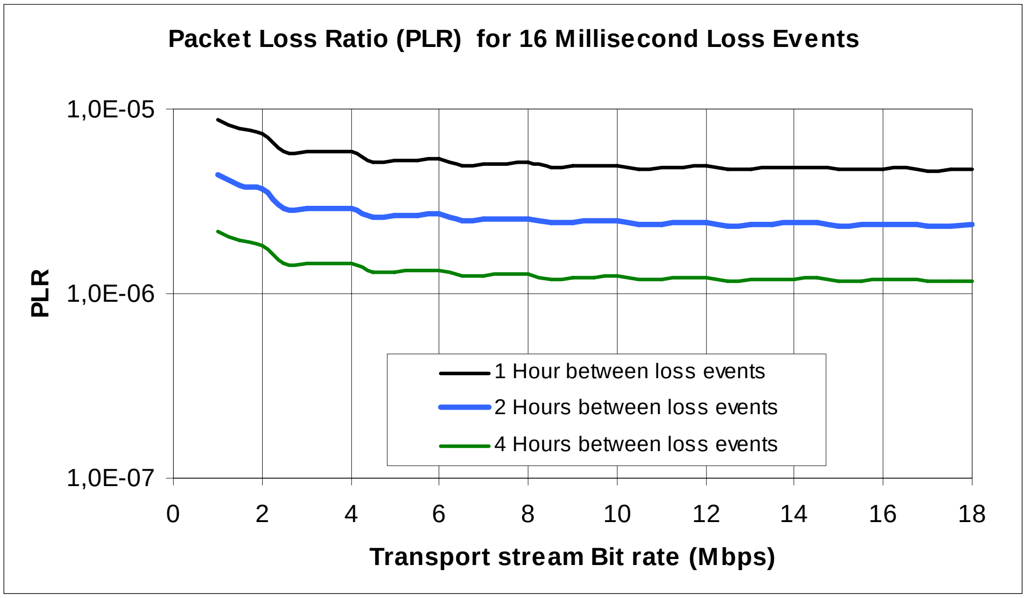


Figure 14: PLR required to meet average time between loss events of 1, 2, and 4 hours assuming each event is an uncorrectable DSL error that loses 16 milliseconds of contiguous data.

Severe Error Limits for SD and HDTV Services

In addition to average packet loss rates impacting picture / audio quality and availability metrics, it may also be advantageous to define a second set of limits on severe impairments. These limits would apply to quality degradations that fall between the impairments generated by the packet loss limits specified above and total service outage (i.e. black screen) metrics specified by the dependability metrics. These gross impairments could include video frame drops, frame repetitions (freeze frames), or short duration (less than 10 seconds) loss of intelligible audio or video or control (e.g. due to protection switching). Metrics are TBD based on industry input and could be specified by frequency of the error event per time unit – e.g., a maximum of one severe error per day and the duration of the impairment.

Appendix III – Gross Error Detection provides a summary of subjective research undertaken on gross error impacts to quality of experience and guidelines for setting limits for gross errors.

7. Voice Quality of Experience Objectives

This section outlines the factors that impact voice services QoE, how voice quality is measured and the network and application layer performance objectives necessary to achieve satisfactory voice service quality of experience in triple-play deployments. As in the video service quality section, performance targets (objective metrics and targets/limits) are provided but this section does not provide a comprehensive design guide. Additional details on voice service requirements and design practices may be found in the Voice Guidelines section of the IV Bibliography.

7.1 Voice QoE Dimensions

This section summarizes dimensions contributing to QoE for conversational voice in a Triple-play implementation. At the voice Service Layer, quality of experience dimensions include:

- Control Plane:
  - Interactive responsiveness (call set-up, control and teardown)
- Data Plane:
  - Voice Intelligibility
    - Many potential impact points on voice intelligibility in an end-to-end system
    - Distortion, loss, delay, echo, and transcoding are common impairments
- Usability
  - Service UI (set-up, directories, caller ID, configuration, etc.…)
- Reliability / Availability
- Security / Privacy
  - for user, and Telco
  - security impacts on other dimensions (ex. encryption / decryption delay)

A general overview of VoIP customer premises deployment options may be found in TR-110 DSLHome™ Reference Models for VoIP Configurations in the DSL Home[53]. The voice service QoE guidelines that follow apply to both voice sessions from pure VoIP phones connected directly to the IP network and from traditional analog phones sitting behind terminal adaptors (ATA).

There are four key factors affecting QoE for VoIP:

- delay (including delay variation or jitter),
- the speech codec,
- cell/packet loss,
- echo.

These factors are included in the ITU G.107 E-model ***R*** measurement for predicting conversational speech quality as shown in Figure 15. A fifth factor: signal level, is not affected by IP transport, but it is important to establish proper settings at the point where an IP network connects to another type of network.

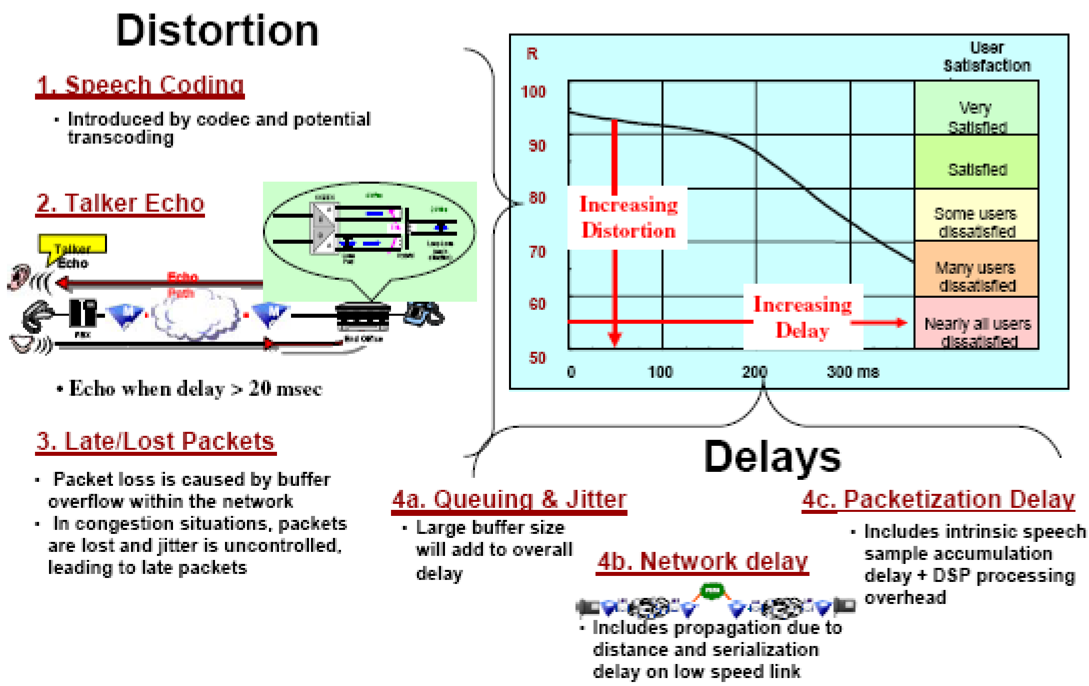


Figure 15 Summary of the voice QoE impairments and the impact on the E-Model "*R*"

Speech codec

The speech codec chosen will have a strong influence on the final obtained quality, both because of the baseline quality of the codec (that is, the quality of the codec without other impairments) as well as the response of the codec to other factors, such as presence of background noise, packet loss, and transcoding with itself or another codec. The choice of codec is an important determinant of the overall performance of VoIP.

End-to-end delay

The end-to-end delay of a voice signal is the time taken for the sound to enter the transmitter at one end of the call, be encoded into a digital signal, travel through the network, and be regenerated by the receiver at the other end. Delay is sometimes called latency. When delay is too long, it may cause disruptions in conversation dynamics. As well, increasing delay makes echo more noticeable.

Jitter

Variation in delay, caused by differences in the time taken for packets to cross the network, is called jitter. Jitter is a concern because the decoding of the digital signal is a synchronous process and must proceed at the same constant pace that was used during encoding. The data must be fed to the decoder at a constant rate. Variation in packet arrival times is smoothed out by the jitter buffer, which adds to the end-to-end (mouth-to-ear) delay. Jitter is not considered a separate impairment because the effects of jitter in the packet network are realized in the output either as delay (added as the jitter buffer wait time) or as distortion from packet loss (because packets arriving after the expiration of the jitter buffer wait time are not included in the output signal).

Packet loss
As in other packet-based services, packets may be dropped during their journey across the network. Packets may also be lost if they are late in arriving at the destination codec buffer and miss their turn to be played out. The missing information degrades the voice quality, and a Packet Loss Concealment (PLC) algorithm may be needed to smooth over the gaps in the signal.

Echo control
Because of the longer delay introduced by VoIP (compared to POTs or TDM voice networks), echo control is a major concern. A given level of echo sounds much worse when the delay is longer. Echo control at the appropriate places in the connection will protect the users at both ends. Echo control relies on the correct signal levels (see Signal Level, below) as well as on echo cancellers and other techniques that prevent or remove echo from the connection.

Signal level
The level or amplitude of the transmitted speech signal is determined by amplitude gains and loss across the network. There are a number of contributors to the final signal level, and most are defined in the loss plan (sometimes called the loss/level plan) of the network. The loss plan for TDM ensures that the output speech is heard at the proper level and contributes to the control of echo. The loss plan for VoIP is reasonably simple; the sensitivities of the sending device (say, an IP phone) and the receiving device (say, a media gateway) are defined by standards, and there is no gain or loss in the packet portion of the network.

Things are more complicated when a packet network is connected to another network with a different loss plan. When the other network is a traditional network with analog access, it may be necessary to adjust the level of each signal path (the signal sent to the other network and the signal coming from the other network) to account for the loss plan of that network. The required loss for each path must be determined and set accordingly. Errors in the loss settings can cause incorrect speech level or audible echo at one or both ends of a connection.

7.2 Voice QoE Measurement

Each of the QoE contributing factors discussed above (and other, conventional voice impairments, such as noise and harmonic distortion) can be measured individually. However, it is useful to have an overall indicator of voice quality. Various metrics have been devised to quantify the overall perceived voice quality of a component or a system. Three common metrics are discussed below:

- the subjective measure called Mean Opinion Score (MOS),
- an objective MOS estimator called PESQ (Perceptual Evaluation of Speech Quality, pronounced “pesk”),
- a computed metric called Transmission Rating (***R***), which is calculated from objective measurements of fifteen contributing parameters using an ITU G.107 standard tool called the E Model .

The quality of a voice call is determined by the access types, the transport technology, the number of nodes the call passes through, the distance, packet transport links speeds, and many other factors that differ from one connection to another. To compare networks, specific connections (reference connections) representing equivalent calling conditions are defined that can be measured and compared.

Types of Mean Opinion Score (MOS)

Mean Opinion Score began life as a subjective measure. Currently, it is more often used to refer to one or another objective approximation of subjective MOS. Although all “MOS” metrics are intended to quantify QoE performance and they all look very similar (values between one and five with one or two decimal places), the various metrics are not directly comparable to one another. This can result in a fair amount of confusion, since the particular metric used is almost never reported when “MOS” values are cited. ITU Rec. P.800.1 provides more details on the distinction between different types of MOS, and how to distinguish them. There are fundamental differences between individual metrics, and numerical values are not necessarily directly comparable just because they are both called MOS.

Subjective MOS

Subjective MOS is a direct measure of user perception of voice quality (or some other quality of interest), and is thus a direct measure of QoE. Subjective MOS is the mean (average) of ratings assigned by subjects to a specific test case using methods described in ITU-T P.800 and P.830. Subjective MOS can be obtained from listening tests (where people rate the quality of recorded samples) or conversation tests (where people rate the quality of experimental connections). Quality ratings are judged against a five-point scale: Excellent (five), Good (four), Fair (three), Poor (two), and Bad (one). MOS is computed by averaging all the ratings given to each test case, and it falls somewhere between one and five. Higher MOS reflects better perceived quality.

Mean Opinion Scores (MOS) are not a measure of acceptability. While perceived quality contributes to acceptability, so do many other factors such as cost and availability of alternative service. Subjective MOS is strongly affected by the context of the experiment (ex. the order in which the test cases are presented in the experiment, the range of quality between the worst and best test cases used in the experiment, and whether the subjects are asked to do a task before making a rating). There is no “correct” subjective MOS for any test case, process, or connection. This is extremely inconvenient, since it means that it is not possible to specify performance or verify conformance to design specifications based on subjective MOS, but it is very important to take account of this in any analysis or design decision relying on subjective MOS evaluation.

PESQ (P.862)

Subjective studies take significant time and effort to carry out. MOS estimators such as PESQ (Perceptual Evaluation of Speech Quality) can provide a quick, repeatable estimate of distortion in the signal. However, the score does not reflect the conversational voice quality, since listening level, delay, and echo are excluded from the computation. Separate measures of these characteristics must be considered along with a PESQ score to appreciate the overall performance of a channel.

P.862 is an intrusive test, which means that the tester must commandeer a channel and put a test signal through it. To perform a test, one or more speech samples are put through a device or channel, and the output (test signal) is compared to the input (reference signal). The more similar the two waveforms, the less distortion there is, and the better the assigned score. The algorithm does some preprocessing to equalize the levels, time align the signals, and remove any time slips (where some time has been inserted or deleted). PESQ then applies perceptual and cognitive models that represent an average listener's auditory and judgment processes.

A diagram of the process is shown in **Figure 16**. The raw PESQ score is usually converted to a MOS estimate using one of several available conversion rules, for example, PESQ-LQ.

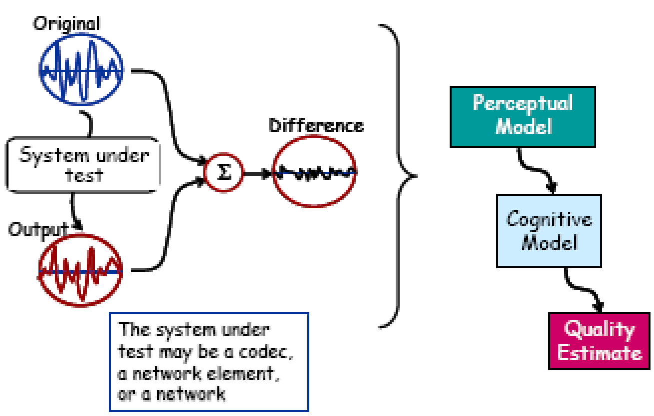


Figure 16 Block diagram showing the operation of PESQ and similar objective speech quality algorithms.

As shown in **Figure 16**, a known signal is input to the system under test, and the algorithm analyzes the difference between them by applying a model of human auditory perception followed by a model of human judgment of preference to arrive at a quality estimate.

Note that many objective quality algorithms have been defined. Aside from PESQ, the best-known are PSQM (Perceptual Speech Quality Measure), standardized as P.861, and

PAMS (Perceptual Analysis Measurement System), a proprietary method developed by BT. As the current standard, P.862 is preferred to the older measures.

Note: MOS estimates determined by different methods are not comparable. A PESQ score is not comparable to a PAMS score, nor is either of them comparable to a score obtained from a subjective test. The E-Model (below) can also be used to compute a "MOS" estimate. MOS computed with the E-Model is **not** comparable to MOS obtained from other methods.

Transmission rating (*R*)
Transmission Rating or ***R*** is an objective metric indicating the overall quality of narrow-band conversational voice. ***R*** is the main output variable of the ITU E-Model (Rec. G.107). Fifteen parameters are used to compute ***R***, some of which are listening level, noise, distortion, the codecs used, packet loss, delay, and echo. Because ***R*** accounts for all the factors that contribute to the conversational voice quality, it is the only value needed to completely describe the quality. ***R*** can be determined for voice calls on any technology platform or combination of platforms (analog, digital, TDM, ATM, or IP) and with any type of access (analog loop, digital loop, wireless, or 802.11).

The E-Model computes ***R*** for individual connections. By defining many connections of interest (called hypothetical reference connections or HRX), we can determine ***R*** for many paths through the network. Generating ***R*** for a well-chosen set of HRX gives a good indication of the performance of a network. By comparing similar call scenarios for different networks, we can get a picture of where a network does well and where it may disappoint users.

The input values used to compute ***R*** can be measured values or expected values. This means that the E-Model can be used to predict the quality of equipment and networks that are still in the planning stages. It is also helpful to compute ***R*** for a benchmark network, particularly where the benchmark provides a "known user experience. Useful benchmarks include:

- The TDM PSTN (Public Switched Telephone Network)
- an existing network that is being replaced by the new network
- a network that delivers a known user experience can be chosen to serve as a quality target, for instance, an ordinary wireless cellular network.

R values can be tabulated to facilitate comparisons, however, it is often useful to make comparisons graphically. E-Model output can be used to generate a graph showing how ***R*** changes as delay increases. The ***R*** x delay relation not only indicates how a particular scenario will respond to increasing distance between the endpoints, but can also suggest the benefit associated with changes to the end-to-end delay. The range for delay is 0–500 ms, which goes slightly beyond the limits suggested in G.114.

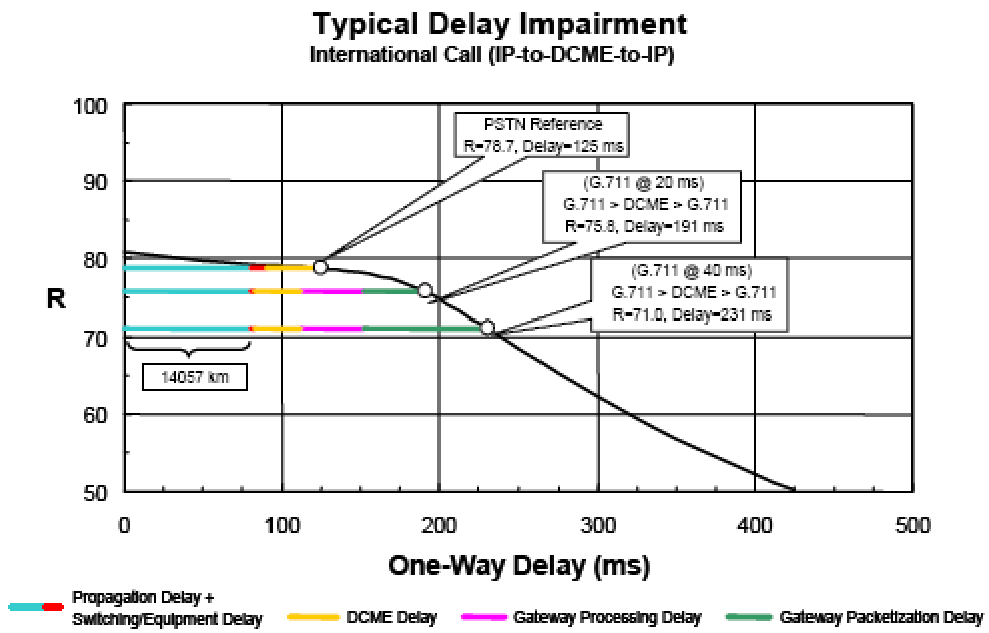


Figure 17 *R* vs. delay for a particular class of terrestrial international calls.

In Figure 17, G.711 is used in the national links with a Digital Circuit Multiplexing Equipment (DCME) link, which generally uses G.726 speech coding at 32kb/s, in the undersea cable. Specific points on the curve show *R* for the benchmark (PSTN reference, TDM end-to-end), and for each of two calls using IP in the national portions of the call (20-ms and 40-ms packets, respectively). Bars under the curve indicate the sources for the cumulative delay associated with each call. Since only one coding scenario is considered (G.711> G.726 > G.711), the model generates only one contour. The model assumes best practices for any factors not specified.

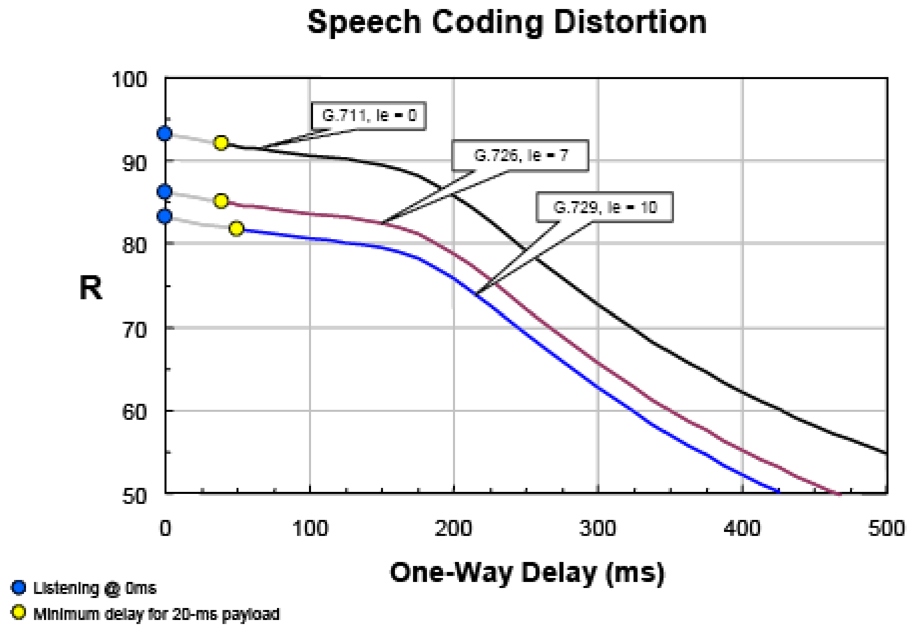


Figure 18 *R* vs. delay for G.711, G.726 (32 kb/s), and G.729 (8 kb/s).

In Figure 18, the model assumes best practices for any factors not specified. Note that although *R* is plotted for all delays, there will be a non–zero minimum delay (yellow points) for interactive calls. For these points, propagation delay is zero. This is the lowest delay for the modeled call scenario (the minimum delay will depend on the codec as well as the packetization selected). In this chart, we have assumed similar equipment delays beyond those associated with the codec; however, in actual network situations, these can change as well. The blue points represent the quality differences heard when listening to recorded speech samples.

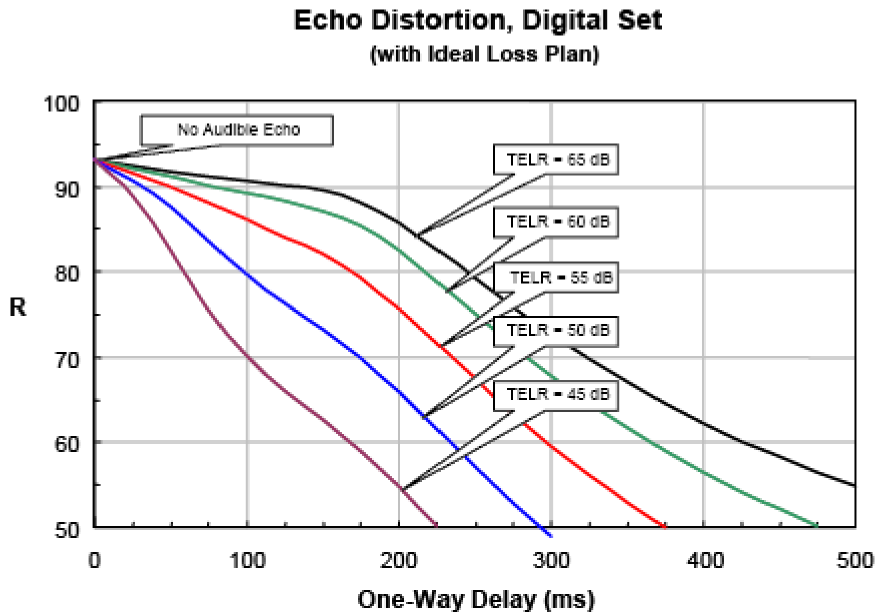


Figure 19 *R* vs. delay for various levels of echo.

In Figure 19, note how *R* drops off more quickly with smaller values of TELR. The increasing rate of degradation for louder echo reflects the interaction of delay and echo discussed above. The model assumes best practices for any factors not specified.

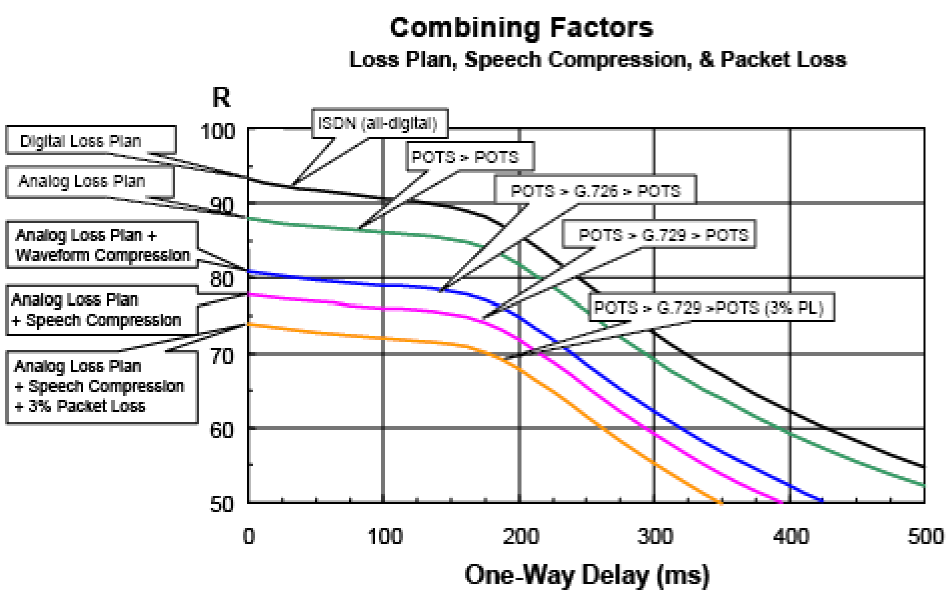


Figure 20 *R* vs. delay for multiple distortion factors

Figure 20 illustrates the effect of successive addition of non–ideal factors: loss plan, compression coding, and packet loss. Since delay does not exacerbate any of these factors, each contour has the same relative shape as the one above. The model assumes best practices for any factors not specified.

7.3 Voice QoE Requirements Guidelines

QoE targets for voice services can be based upon absolute threshold or can be set relative to a known user experience. The latter one is the one preferred for voice services since end-users have long time experience with telephone systems. Note that defining voice QoE targets is not a trivial task and requires the appropriate subjective evaluation expertise along with human factors behavior knowledge. An approach based on known user experience implies that there might be multiple targets, and no single target is applicable to all situations. For example, mobile users have different expectations than wireline users. Similarly, people making international / overseas calls have different expectations than on local calls. Therefore, there might be multiple targets required depending on the call scenarios supported by the network.

Based on the ITU E model and R transmission quality rating, it has been determined that a difference of 3***R*** is not noticeable by typical users and, therefore, a triple-play service offering should be engineered within this margin in order to provide an equivalent replacement technology (ex. PSTN). Differences of 3-7***R*** might be noticeable, but most likely acceptable. Larger ***R*** degradations (greater than 7***R***) are more likely to be noticeable and should be avoided. Table 16 lists the recommended end-to-end QoE voice service performance objectives in triple-play service deployments.

Services	User QoE Performance Targets	
	QoE metrics	QoE targets
Conversational voice (CBR and VBR)	Conversational Voice	
	R-factor	ΔR PSTN - Packet < 3R
	delay	< 150 ms
	distortion	Ie < 3R
	Path Interruptions Due to Failure	
	Frequent Interruption	80ms (affects speech intelligibility)
	Infrequent interruption	3 sec (perceived as call drop)

Table 16 Voice QoE Requirements Guidelines

Note the QoE performance targets listed in Table 16 are for the complete end-to-end voice call. Since an end-to-end voice call will typically traverse multiple networks, the impairment objectives shown in Table 16 will need to be distributed across all the networks and cannot normally be consumed by a single network. ITU-T Recommendation Y.1541 provides guidance in this area.

8. Best-Effort Internet Access QoE Objectives

Currently, broadband access is used mainly for access to the Internet and there are no guarantees on the quality of service of the transport layer as the Internet itself is based on best-effort transport. Since the transport service is best effort, and services could be provided from outside the transport service provider's control, no quality of service guarantees can be made. However, target QoE metrics and values can be established for these applications and services in order to satisfy user expectations.

The primary best-effort applications are web browsing, e-commerce, email, instant messaging (IM), and file transfer. The presentation of the data may be audio-visual, graphical or textual. Real-time, interactive applications such as VoIP, gaming and media streaming are becoming popular over the best-effort Internet network as well. ITU-T Recommendation G.1010 "End-user multimedia QoS categories"[9] provides additional details on the taxonomy of applications.

Strictly QoE should be defined for each application type but G.1010 provides a broader classification of applications with common requirements ranges as shown in Figure 21. Applications fall into 4 categories according to the interaction delay requirements for satisfactory QoE and whether they are error tolerant. As noted in the previous sections, although audio and video applications can tolerate some errors / loss, the QoE can be impacted if the loss is too large.

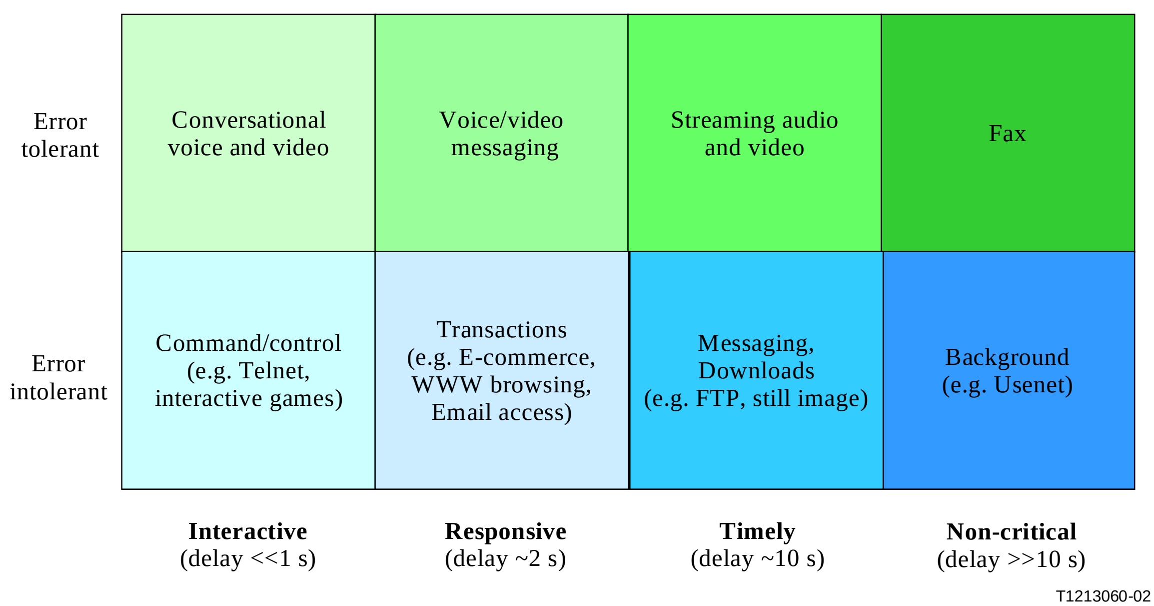


Figure 21 ITU G.1010 Application Classes[9]

As shown in Figure 21, applications such as web browsing, e-commerce, IM/email access fall into the Responsive category and gaming in the Interactive category. Underlying all the Responsive applications are QoE parameters that are related to the fundamental service features which are typified by web browsing. Audio and video QoE parameters are discussed in greater detail in previous sections and could be used as targets in a well engineered best effort access scenario as well.

The purpose of this section is to identify the QoE parameters appropriate to Web browsing as a good representative of Responsive applications, to determine the way in which they are related to network performance and service quality parameters in order to ensure a satisfactory user experience for applications in the Responsive category. Some additional details for gaming applications typifying the Response category will also be provided.

The ITU guidelines in G.1010 have been taken as a starting point for performance objectives, but as the QoE requirements for specific applications are understood in more detail, the guidelines could be revised if this would lead to a significantly better QoE for the end-user. These and other services may be migrated to a premium offering with guarantees of network performance that would also ensure QoE requirements. As the applications are migrated beyond best effort delivery, QoE requirements can be examined in greater detail (as done in prior sections of this document for Entertainment Video and Voice applications) and targets revised as required.

8.1 Best-Effort Web Browsing QoE Dimensions

As with voice and video discussed previously, the dimensions impacting service quality occur at both the network layer and at the application layer commonly on the end-points, e.g. the characteristics of the application server and/or the end-user terminal. The fundamental quality of experience dimensions for best-effort Internet applications typified by web browsing are as follows.

- **Initial system response time** (e.g. delay) from providing URL to the end-user being aware that a download has started
- **Data download speed**. This is frequently communicated to the end-user by means of a file transfer dialogue box, numeric display of download status (% or number of bytes downloaded / total bytes) or by a network rate meter of some kind. For smaller downloads, the speed can be indicated by the rate at which the screen updates
- **Consistency of download speed**. If the download is at a steady rate, then the user has a good idea (either intuitively or through a meter) of when it will finish. If on the other hand the rate varies greatly, then the finish time is much less certain and the user cannot plan how to use the intervening period so effectively. Note that the underlying bit error rate or congestion induced packet loss will have an impact on the consistency of download speed.

- **Incremental Display –** the time before there is some new, intelligible content to view. It is often the case that the display starts to update before the download is complete. As soon as the user has some new information to consider, the fact that the download is still in progress is of less importance.
- **Action availability –** the time before the user can undertake the next step in the browsing process e.g. a new link or action button becomes useable; again this may be before the download is complete.
- **Time until the download is complete**. In addition to the download rate, the user experience is also impacted by the absolute time that a download has taken
- **Usability**
  - Application User Interface (initial set up speed, navigational aids)
- **Content**
  - information quality/quantity
  - information presentation
- **Availability** of the content source.
- **Security/Privacy**
  - for user, service and network providers and content owners
  - security impacts on other dimensions (ex. encryption/decryption delay)

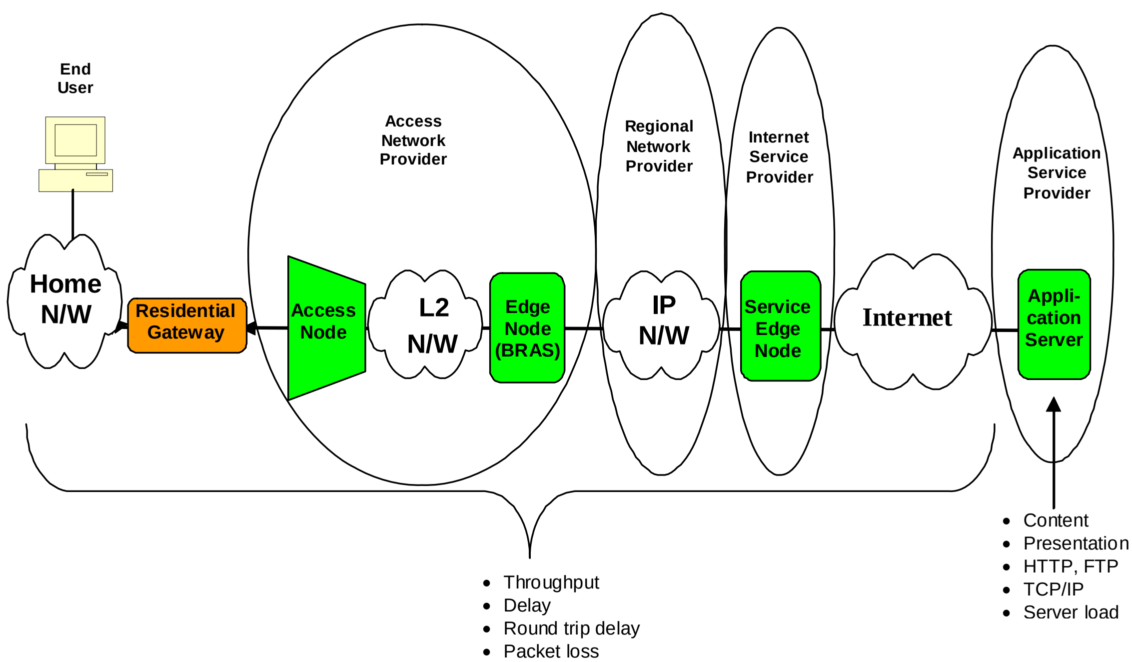


Figure 22 End-to-End Delivery of Best-effort Internet Access

Figure 22 shows the network elements and networks involved in an end-to-end best-effort Internet access connection and the various providers involved. It also indicates the key factors affecting QoE. In the transport network the QoS parameters of importance are throughput, delay and packet loss. Packet loss does not result in data errors because best-effort internet access uses TCP, but it will affect throughput, add delay and increase rate variability.

8.1.1 Network Layer - Factors Impacting Download Rate

It is increasingly common for service providers to compete on the basis of peak access network rate offered. However, there are several reasons why the advertised rate may not always be seen by the end-user in practice when running a BE application. Note that most of the other performance limits on the bandwidth available for an application have not yet been reached, but they do indicate that beyond a certain limit there is little point in increasing the access rate alone.

The other application bandwidth performance limiting factors are:

- While TCP provides reliable transport (via retransmission) it can also limit the download rate. In general, the TCP throughput depends on the following parameters: Round Trip Time (RTT), maximum window size (normally up to 65536 bytes), packet loss ratio and Maximum Segment Size (MSS) (typically set at 536 bytes but could go up to 1480 in Ethernet networks). TCP uses a sliding window approach, with the window size being the maximum amount of data that can be sent before an acknowledgement (ACK) must be received by the sender. The throughput of the TCP RENO algorithm can be determined using the following formula: min{max_window/RTT , MSS/(RTT*sqrt(loss))}

  In case the packet loss ratio is sufficiently low (say lower than 10-5), then this can be simplified to the window size divided by RTT.

  Note also that a certain amount of upstream bandwidth is also needed for the ACKs, a good rule of thumb is 10% of the real downstream data-rate. If this amount is not available, either because of a highly asymmetric access system, or the demands of higher priority applications e.g. VoIP, then the download rate will be reduced.

- The network operator / service provider needs to provide aggregate capacity, and while this capacity must be able to support the peak access rate for at least some number of users (in the limit, one), it is commercially difficult to provide the full capacity required to aggregate all users at their peak access data rates, particularly if the higher peak access rates do not provide additional revenue.

- Many information sources are accessed via the Internet, i.e. at least part of the Network is beyond the control of the Network Service Provider. Constrained interconnect bandwidth at Internet peering points can for example become a performance bottleneck.

- Home networking technology. As access rates approach 10 Mbps, some Home Networking technologies, notably some Wireless and Power-line systems, can become the performance bottleneck. Note also that such technologies are inherently time-varying at the physical layer (due to time varying noise and interference) and so can also be responsible for download rate variability

- The PC itself. The PC can be a performance bottleneck in pure network speed terms, in addition to the impact of other applications (see below).

The other issue is of course the degree to which the end-user actually notices an improvement in performance, even when there is one. End-users are very unlikely to pay for an increase in performance that they do not notice.

8.1.2 Application Layer QoE Factors

Application layer factors impacting service quality are as follows:

- Content issues - the quality, attractiveness, level of interest, importance, etc. of the content itself.
- The speed of response of the server. This will be influenced by the capabilities of the server to service the quantity of contemporaneous download requests it is receiving.
- The speed of response of the DNS server.
- The end-user terminal capabilities will affect the receiving and displaying of data.
- The number of other applications running on the end-user terminal.

Content impacting issues are beyond the scope of this work. Any end-to-end performance objectives provided should however take into account the entire system from server to "well-behaved" browser. Misconfigured applications, abnormally slow PCs, etc. are beyond the control of the network service provider in most cases.

8.2 Best-effort Internet Access QoE Measurement

The responsiveness component of QoE for best-effort internet access can be measured in three ways:

1. Subjectively – using a controlled usage experiment and experiment participants who grade the quality using rating scales such as Mean Opinion Score (MOS).
2. Objectively – using electronic test functionality to measure various aspects of the downloading experience such as web page load delay
3. Indirectly – using measurements of network performance parameters (throughput, delay) to estimate the impact on the downloading experience.

Subjective experiments are time consuming and costly. There is some published work on the relationship between user perception and computer response times over a wide variety of application types as shown below:.

- People prefer response times < 1 sec. Although people can adapt to longer times, they are generally dissatisfied: Schneiderman (1998)[54]
- Delays must not exceed 1 sec for uninterrupted thought: Nielsen (1994) [34]
- Users become aware of waiting if response delays exceed 2s; impacts include: decreased productivity, increased frustration levels, short term memory deterioration: Gallaway (1981)[55], Miller (1968)[56], Youmans (1983)[57], Williams (1973)[58]
- Number of productive transactions drops when system response time exceeds 4s, and the drop is dramatic beyond 12s: Barber & Lucas (1983)[59]
- For web browsing - after 8.5s users assume the system has hung/broken unless feedback is provided: Bickford (1998)[60]
- Response delay must be less than 10 sec to maintain attention: Nielsen (1994) [34]

As illustrated by the examples above there appears to be a threshold of acceptable delay in the 2-4 second range and a time of 8-10 seconds beyond which users will give up, for a variety of applications and interactive tasks tested including web browsing. Beerends and van der Gaast (2004)[61] performed web browsing specific quality modeling and found that in general, quality perception in relation to response time fall within three broad perceptual regions that can be distinguished as follows:

- Instantaneous experience: 0.1s is about the limit for having the feeling that the system is reacting instantaneously, an important limit for conversational services (e.g. chatting).

- Uninterrupted experience: 1.0s is about the limit for the user’s flow of thought to stay uninterrupted, even though the user does lose the feeling that the service is operating directly, an important limit for interactive services (e.g. web browsing).

- Focused experience: 10s is about the limit for keeping the user’s attention focused on the dialogue. For longer delays, users will want to perform other tasks while waiting for the computer to finish, so they should be given feedback indicating when the computer expects to be done. Feedback during the delay is especially important if the response time is likely to be highly variable, since otherwise users will then not know what to expect.

For objective measurements the observable parameters are measured at the user terminal. The web browsing application delay impacting QoE includes:

1. Delay between entering a URL and indication that a website has been found
2. Delay to first indication of the start of a download
3. Delay to reception of sufficient data for the user to start absorbing information

4. Delay to reception of sufficient data for the user to start using the data
5. Delay to completion of the download
6. Percentage of download cancellations (in relation to Bandwidth and/or Object Size)
7. Response delay variation for interactive sessions (such as gaming).

Parameters 1, 2, 5 and 6 are easily measurable on test data and live data. 3 and 4 could be measurable by careful design of test pages but for live data would require some mechanism to detect when data begins to make sense to the user and when it becomes useful. Parameter 7 can be measured in and out of service.

Using measurements of network quality of service parameters to indirectly measure QoE is problematic. The two network QoS parameters of interest are delay and throughput. There are two layers that impact the latency experienced by a user:

- Network layer – packet delivery latency
- Application layer – server latency (e.g. responding to user requests, generating data)
- Application layer - client latency (e.g. rendering images).

The end-to-end delay and throughput perceived by the user can be strongly affected by the application layer performance of the server and the user terminal. In addition home networks supporting the user terminal can add delay and reduce throughput and may be beyond the control of the service provider.

However, it is possible to measure round trip delay and throughput from the server to the user terminal by exploiting the fact that TCP is used. Simple additions at the user terminal can enable measurement of the response time (round trip delay including network and server) and throughput from the server. In a live situation if they are outside acceptable limits then it is an indication that the user would be getting a poor QoE. If they are inside limits then this does not mean that the QoE is acceptable because the user terminal performance (e.g. display rendering) may still be poor.

8.3 Gaming Specific QoE Dimensions

Interactive gaming applications have similar key performance dimensions as those that drive satisfactory QoE for web browsing outlined above – namely delay. In the case of gaming, however, the user action (shooting, steering, moving through a virtual space, etc.) generate feedback (gun blast, car moves, character moves, etc.) that is perceived to be immediate. The end-to-end system response time (SRT) is the key gaming QoE parameter and consists of the time the system needs to detect and process the a user-initiated event (e.g. button push on game controller), transport it over a network to a game server, to process it, and to send the updated game state back to the local output device (e.g. PC or TV screen).

8.3.1 Classification of Game Genres

To discuss QoE of interactive Games, it is useful to first classify the most popular (networked) games and discuss which of them are to be considered.

The following covers most of the popular genres being used today:

- Platformer (e.g. Mario Bros™)
- Adventure (e.g. Monkey Island™)
- Role Playing Game (RPG) (e.g. Sims™)
- First Person Shooter (FPS) (e.g. Quake™)
- Third Person shooter (e.g. Tomb Raider™)
- Shoot 'em up (e.g. Commando™)
- Puzzle (e.g. Tetris™)
- Simulation (sports/racing/pilot) (e.g. Collin McRay Rally™, Need For Speed™)
- Turn-based games (e.g. Chess)
- Real Time Strategy (RTS). (e.g. Warcraft™)

Note gambling games, like “turn based” (e.g. Cards) or “time limited” (e.g. betting) games are not considered in this document.

Many of above genres have multiplayer capabilities that range historically from
- two users sharing the same client (e.g. hot seat, keyboard sharing) to
- two users with point2point connection (e.g. serial link) to
- two or more users with a LAN connection (still popular, e.g. LAN parties) to
- two or more users with Internet connection to
- many users with Internet connection.

Initially, online games were simple text based such as the original Multi-User Dungeon (MDU). With time more and more complex graphics have been incorporated, making specially the RPG, FPS and RTS genres popular. Nowadays, Massively Multiplayer Online (MMOs) sub-genres of these ones have become especially popular as a consequence of broadband access adoption, making possible virtual worlds, populated by many players simultaneously (MMORPG, MMOFPS and MMORTS).

8.3.2 Multiplayer Online Games Communication Architectures

It is also important to understand the communication architectures that are used by games. First of all, there are two different paradigms: peer-to-peer and client-to-server.

In the peer-to-peer paradigm, all the nodes (e.g. server, game client) correspond to game instantiations (one for each player) where no node is more special than other. They must be (logically) connected to each other. There are no intermediate nodes and hence each node must take care of communicating with all the rest . The problem of this architecture is that, in general, it does not scale well with the number of nodes, as the number of communication relations grow in a quadratic way. The usage of broadcast or multicast techniques could alleviate this. A benefit of this kind of architecture is that is relatively easy to implement by game coders once that they produced the single player version of the game. Another potential benefit is that latency should be low, as communication is directly performed between nodes. However, the used bandwidth will grow with the number of nodes (even although using broadcast or multicast techniques, as these techniques do not provide a benefit in the downstream access segment), and congestion can happen, increasing delay.

In the client-to-server paradigm, some nodes are promoted to the role of a server. Server nodes act as mediators and arbitrators, reducing the number of communication relations. In the case of having just one only server, the architecture is said to be centralized. In the case of multiple servers, the servers can communicate in a peer-to-peer basis or in a hierarchical way. Note that the server role can be as previously defined, i.e. having a special node running specialized software different from the client software which is distributed to the players, or it can be dynamically allocated between the client nodes (or a subset of them), either manually or triggered by pre-defined conditions. In this way, the software distributed to the players is able to work both as a client and as a server. In this last case, these server nodes are known in the p2p world as super-nodes. Some implementations do not allow a client node to prevent itself from becoming a super node.

In a client-server architectures, communication does not flow directly between all the nodes, and hence the communication delay is expected to be longer. However, the total volume of exchanged information can be lower as broadcast/multicast or 'management of interest' (see the following section) techniques can be used also by the server. Be that as it may, in client-to-server architectures, the server node is a critical part, as it must be able to cope with a variable number of clients located at variable and different positions, whereas the client node does not care about the communication with other clients.

It is possible to find commercial games that can work in several of the previous modes, allowing games to be played among a reduced set of players belonging to a closed community (normally connected to the same LAN) in a peer-to-peer basis or in a client-to-node basis with no predefined server, or to connect to a known server in order to play with players of an open community, especially in the case of MMO.

Typically, MMOs use the client-to-server architecture, especially because in this way, the game play can be controlled by the game developers. Control over the game environment is particularly important since in many of these MMOs with virtual worlds, there exists a real world economy for the virtual world goods where weapons, powers, game characters, etc. are bought and sold for real currency (e.g. Everquest and Sony Station Exchange).

8.3.3 Interactive Games QoE Dimensions

The entertainment value a game provides is influenced by the QoE delivered by the gaming service. It is noteworthy that there is no accepted model for evaluating player enjoyment of games. However there are some proposals as such as GameFlow proposed by Sweetser and Wyeth (2005)[62] where a model based on eight different criteria is proposed. The criteria are the following ones:

- Concentration. Games should require concentration and the player should be able to concentrate on the game.
- Challenge. Games should be sufficiently challenging and match the player's skill level
- Player Skills. Games must support player skill development and mastery - players should be able to start playing the game without reading the manual
- Control. Players should feel a sense of control over their actions in the game.
- Clear Goals. Games should provide the player with clear goals at appropriate times.
- Feedback. Players must receive appropriate feedback at appropriate times.
- Immersion. Players should experience deep but effortless involvement in the game.
- Social Interaction. Games should support and create opportunities for social interaction.

Computer games are typically reviewed over a number of performance dimensions that are selected to indicate the important aspects of the quality of experience when playing the game. A non comprehensive list of aspects that appear in magazine and online reviews of games is the following:

- **Graphics**. Graphical aspects are especially considered. Good graphics, according to the epoch when the game title is launched, act as an attraction to potential purchasers. The graphics quality standard grows every day, and 2D and 3D graphics are expected to be rather spectacular, especially for computer and console games. Smooth movement and good definition are essential aspects. With regards to 3D graphics, the rendering technology used is especially crucial, as it must provide good quality in terms of image definition and at the same time allow fast drawing of graphics, increasing the number of frames per second that can be drawn. The style of the graphics is also important. A 'cool' style, according to the epoch standard, is well appreciated. Status icons that indicate useful information to each player must be also cared for. Good graphics help the user to "concentrate" on the game and hence to "immerse" into it, and also to receive "feedback".
- **Sound**. Sound effects and music are aspects that also appear in reviews. Sound effects are used in many games not only for providing more realism but also provide important information, serving for instance to acknowledge actions that have been commanded by the player or announcing to events that are about to happen. Music helps to give a good atmosphere to the game. Games usually come with a rich set of music themes that are played randomly and/or according to the game status. Speech is also commonly used for communicating information to the player in a more natural and agile way. Besides, it is very common to find in today's multiplayer online games that players can talk with each other in order to interact more naturally and without the necessity of stopping the flow of the game. Hence, sound helps, with "concentration", "immersion" and "feedback". It is also important to note that graphics and sounds must be consistent and well synchronized.

- **Playability**. In Jarvinen et.al. (2002)[63], there is a distinction between functional and structural playability. Functional playability deals with the manageability of the game; that is, the control. A friendly and intuitive control interface is a must. Structural playability is related to the rules of the game. A trade-off between functional and structural playability is especially important so that to control the game does not become an impossible task, but at the same time keeping the game world illusion interesting and believable enough to be played (Kücklich, (tbd))[64]. Hence, the skills of the player and the challenges of the game must be carefully balanced. There are different strategies, such as being able to choose different difficulty levels, online tutorials, increasing the game options as the player immerses more and more in the game, etc. That is, although the game can be complex, the learning process must be cared for. Nowadays, embedded tutorials are commonly available on relatively complex games, so that players can easily face the different aspects of the game and learn the rules in an easy way, without the necessity of reading long and complex handbooks. It is also important to have a good balance of short and long term goals so that even with a small set of game skills learnt, a player can start to play and enjoy the game. Addictiveness is the golden design goal of games developers. Hence, playability is closely related to the afore mentioned aspects of “player skills”, “control”, “clear goals”, “feedback” and “immersion”.
- **Difficulty**. This term is normally used not for valuating whether the game control is difficult to manage but for expressing how difficult is to achieve the main goal of the game. If a game is difficult and addictive, a player can enjoy a long time with it, and the purchase seems well worthwhile. This term can also impact the longevity or lastability of the game. Difficulty is specially related to “challenge” and “clear goals”.
- **Originality**. New game genres hardly ever appear. Because of this, originality normally measures whether the story or narrative of the game is original or whether previously unseen combination of genres has been used. Originality is related to “challenge”.
- **Overall valuation.** Normally, an overall review is provided that tries to balance all the aspects of the game that have been explicitly reviewed and other aspects.

The previous aspects are normally presented with a score that can be used for comparing similar and contemporaneous games. It is also common to find some comments from the reviewers, explaining what aspects were especially liked or disliked.

It is important to remark that the previous list of aspects is valid for computer games in general and, hence, for multiplayer online games in particular. Notwithstanding, some of these aspects can change as a function of whether the game is single or multiple player. For instance, the control interface can be well designed for playing solo but not for playing with multiple players; the movements can be pretty smooth in the single payer version but not in the multiple one.

8.3.4 Network Layer Factors Impacting QoE

Specific aspects for multiplayer online gaming must also be considered. Playability is a factor that can dramatically change in networked games. More precisely any disturbance of the basic illusion expected to be created by the game will decrease the playability of the game: “If the mediated nature of the game experience becomes apparent […], the playability of the system is deficient." Kline et.al. (2003)[65] .

The following aspects of playability are the most critical in networked games:

- Responsiveness. In networked games, this term is related to the delay that it takes for an update event to be registered by the affected players. In other words, the game must react quickly to the inputs of players, rapidly showing their effects. With a networked game, the responsiveness will depend on the network delay experienced.
- **Smooth responsiveness**. In networked games, delay can be variable, and hence responsiveness can vary. It is important to note that players are even able to adapt to a slow “game pace”, developing a mental expectation of the lag between when they issue an order and when the order is effectively executed (Bettner and Terrano, 2001)[66]. That is, a consistent slow response is always better than alternating between fast and slow responsiveness. Hence, a jerky responsiveness is especially bad for playability.
- **Information consistency**. In networked games, consistency refers to the similarity of the view of the game status by the different players. That is, data describing the status of the game must be well synchronized, so that no unfairness is created between players. Absolute consistency means that each player has the same information.

8.3.4.1 Implementation Factors Affecting Network Layer Performance

Design decisions can have dramatic consequences for the QoE of networked games. If one replaces the transfer of all item positions in a simulated world by the exchange of the random seed used as input to the random generator, there can be significant savings in the amount of data that needs to be transported.

Game developers can cope with inconsistency, by ensuring that losing one piece of information is unlikely to make the game inconsistent. E.g. when a single shot can kill a player, then losing this information will immediately lead to inconsistency, where if you need to be shot several times, a single loss will not be a problem.

Interest management is widely used in networked games, and especially in MMO ones, to reduce the amount of information that must be transmitted. The principle is as simple as transmitting just the necessary information to the interested nodes (players). In this way, if two players are not capable of interacting in the game because they are playing in different regions of the “game board/world”, it is not necessary to exchange information between them. Consistency is not affected as they access disjointed parts of the data. So, only the relevant information to each node is transmitted. Obviously there is a trade-off between the complexity of the management of interest areas and the amount of information to transmit.

In client-server architectures, it is easier to benefit from the interest management concept, as the decision of the area of interest is performed by the server/s which has/have a unified view of what is happening in the game. In the peer-to-peer approach, all clients must be informed of the status of the other clients in order to decide whether there is a common area of interest that must be managed.

Finally, another important design decision is related with fair playing and especially with the cheating, which can be very annoying for the players. For this reason, client-to-server architectures are mostly preferred, as the server can perform anti-cheating activities.

8.3.5 Interactive Games QoE Measurement

There is no unified way to measure the QoE, or playability, of networked games. There are some proposals, such as the OPScore (Online Playability Score) model from UbiCom (2005)[67] which operates indirectly using measurements of network performance parameters such as the one-way delay from client to server, jitter and packet loss, and using a G-model (similar to the E-model described in ITU-T G.107[11] for measuring the perceived quality for VoIP applications), try to estimate the impact on the playability of a networked game.

Within the MUSE Project, a G-model for FPS games has been developed (Wattimena et.al. 2006)[68]. In order to obtain the MUSE G-model a large scale subjective experiment was conducted, with a total number of 12 test subjects, playing games of Quake IV, under 33 different network conditions. The subjective experiments lead to the conclusion that ping (i.e. Round Trip Time) and jitter have a negative impact on the MOS, on the acceptability of the game play and on the total number of kills realized within a game session. At the same time, it was shown that for Quake IV packet loss has no noticeable influence on either MOS, acceptability or the total number of kills.

Making use of multi-dimensional regression analysis the MUSE G-model was developed, which enables us to predict a gamer's quality rating (expressed in a MOS) based on measured ping and jitter values. The MUSE G-model exhibits a very high correlation (correlation coefficient R = 0.98) with the subjective MOS values and therefore the model can be used for predicting the gamers quality rating. The MUSE G-model is defined as follows:

$$\text{MOS\_gaming} = -0.00000587\, x^3 + 0.00139\, x^2 - 0.114\, x + 4.37,$$

where x denotes the network impairment and is defined as

$$x = 0.104\ \text{ping\_average} + \text{jitter\_average}.$$

From the form of the network impairment x we conclude that with regard to the impact of Network Performance on interactive gaming quality (at least for FPS games), the most important value is the jitter, followed by the Round Trip Time and the least important factor is the packet loss rate.

With regards to subjective measurements, MOS is commonly used. In fact, in Dick et.al. (2005)[69] a concrete MOS scale is defined:

MOS	Description
1	unacceptable environment impossible to play game
2	very annoying environment server change necessary
3	clearly impaired environment although still acceptable
4	minor impairment noticeable very good environment
5	no noticeable impairments perfect environment

8.3.6 Interactive Games QoE Guidelines

For online games, the network impacts on the QoE are mainly decided by the selection of the traffic class by the network operator. The following figure shows a categorization of traffic classes. The traffic classes are used to discuss what is appropriate to fulfill the interactive gaming requirements:

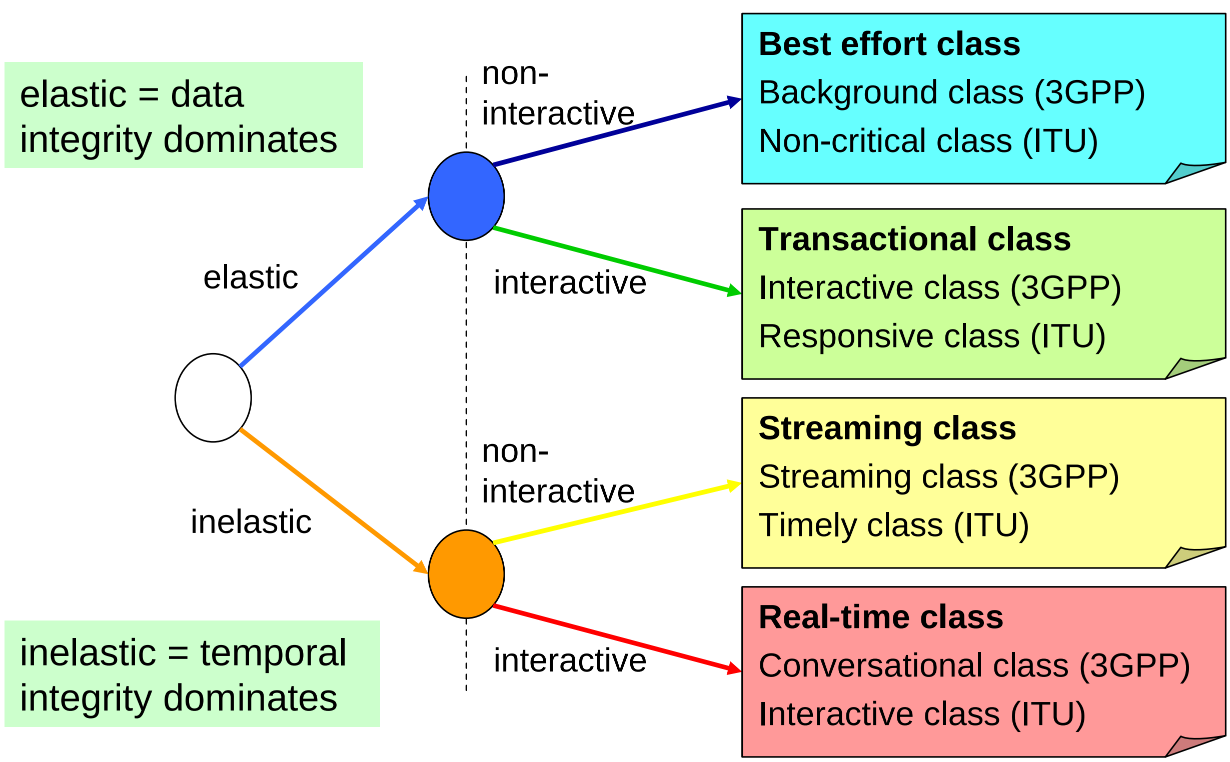


Figure 23 Traffic Class Categories

To achieve the high consistency of interaction required for satisfactory game QoE, it must be guaranteed that game processes running on the remote nodes (where players are located) are tightly coupled (Smed et.al.)[70]. This synchronization essentially requires no loss of information. That is, consistency requires data integrity and therefore associated traffic should have to be classified as elastic and interactive.

To achieve high responsiveness required for satisfactory game QoE, data must be distributed among affected players as quickly as possible, needing a low network delay. Hence, responsiveness requires time integrity, and hence associated traffic should have to be classified as inelastic and interactive.

Obviously, game traffic cannot be elastic and inelastic at the same time, and it is not possible to achieve both high consistency and high responsiveness in a networked architecture (and in particular in a networked game) at the same time. So it is necessary to prioritize one of these QoE targets over the other,.

With regards to the smooth responsiveness, this requires a stable delay, and it is common to sacrifice mean delay in order to guarantee a low delay variance (a.k.a. jitter). Therefore, it is common that different nodes involved in a game try to adapt to the one who suffers the longest delay by means of using buffering techniques (Bettner and Terrano, 2001)[66]. Notwithstanding, this adaptation technique must be used up to a given limit, in order not to make the game unplayable by all.

Responsiveness, or rather a smooth or constant responsiveness is the primary QoE factor that must be achieved. Hence, networked games traffic should belong to the “real time” traffic class, which is inelastic and interactive, and the network should handle the corresponding traffic as for real-time voice and video communications or even better. However in practice, gaming traffic is often carried as best-effort which is the topic of this section. The QoE requirements however are independent of the traffic classification or QoS mechanisms used.

When looking at the three most important genres, FPS games have the highest demand in responsiveness, where RPG have more need for consistency, and RTS games do not need specially high responsiveness, but have consistency demands. This is reflected by the following table:

Game Category	Responsiveness	Consistency
FPS	Highest	Lowest
RPG	Medium	Highest
RTS	Lowest	Medium

Table 17 Game Requirements by Category

To determine specific delay requirements for interactive gaming, there are a number of studies that indicate the delay should be in the range of 100 ms round trip for a user action to screen update. Examples are listed here:

- 100 ms is limit for users to perceive the system is providing “instantaneous” response: Nielson (1994)[34]
- Key press to character display delay should be faster than 100ms with increased typing errors associated with delays beyond 100ms: Long (1976)

There have also been a number of studies on online gaming specifically. Quax et al. (2004)[71] investigated the objective and subjective influence of delay and jitter as present in typical access networks on the quality of game play. They used an experimental setup consisting of a LAN in which 12 players competed against one another in the FPS (first person shooter) game Unreal Tournament (UT) 2003. A router in the network simulated delay and jitter on the network connection for some of the players. Findings indicated that from a user perspective, the players’ perception of the quality of the game depends on the size of the delay the network introduces, with indications that for round trip delays of 60 ms or greater, the players’ perceived the experience to be inferior and judged the delay impairment to be disturbing.

In another study, Beigbeder et al (2004)[72] found that latencies as low as 100 ms can significantly degrade performance for shooting with precision weapons both in terms of accuracy and game responsiveness. As latencies increase above 100 ms, shot accuracy continues to decline further, with a decrease of over 50% at a latency of 300 ms. Occasionally players were also able to notice packet loss when induced loss rates were at least 3%, with the primary artifact noticed being that the game would sometimes not display animations for shots that were fired.

According to Pantel and Wolf (2002)[73] for racing games, delay values over 100ms show significant impact on the realism of the game. Armitage (2001[74], 2003[75]) showed that most of the so-called ’hard-core’ gamers often simply choose not to connect to game servers that show a ping higher than a few 100 ms. Schaefer et al (2002)[76] using the conceptual understanding and theoretical application of MOS tests to assess the quality of multiplayer games found that if a MOS of 3.5 was considered commercial game quality, then a SRT of 139 ms would provide adequate game quality for a multiplayer game such as XBlast.

8.4 Best-effort Internet Access QoE Guidelines Summary

There has been work in the past in the ITU on defining QoE parameters for data transfer applications. These have been simply in terms of delay and error tolerance. Table 18 below is drawn from ITU-T Recommendation G.1010 (Nov 2001) and gives the delay figures for different types of data transfer applications. Unfortunately, it is not clearly stated in G.1010 whether the figures are one way delay from server to user or responses times although this can be inferred from the application. It is clear that both the amounts of data, and the acceptable delays do not accord with current experience and need significant revision[1], however they can be used as the staring point for some basic performance calculations. The web-browsing delays specified in G.1010 for example align well with the other research listed above but the gaming and command / control times listed in G.1010 are likely too high. Response time variation is also an important parameter, but this is not specified in G.1010. Figures for this parameter need to be added where applicable when this table is revised.

Application	Degree of symmetry	Typical amount of data	One-way delay or response time
Web-browsing – HTML	Primarily one-way	~10 KB	Preferred < 2 s /page Acceptable < 4 s /page
Bulk data transfer/retrieval	Primarily one-way	10 KB-10 MB	Preferred < 15 s Acceptable < 60 s
Transaction services – high priority e.g. e-commerce, ATM	Two-way	< 10 KB	Preferred < 2 s Acceptable < 4 s
Command/control	Two-way	~ 1 KB	< 250 ms
Still image	One-way	< 100 KB	Preferred < 15 s Acceptable < 60 s
Interactive games	Two-way	< 1 KB	< 200 ms
Telnet	Two-way (asymmetric)	< 1 KB	< 200 ms
E-mail (server access)	Primarily one-way	< 10 KB	Preferred < 2 s Acceptable < 4 s
E-mail (server to server transfer)	Primarily one-way	< 10 KB	Can be several minutes
Fax ("real-time")	Primarily one-way	~ 10 KB	< 30 s/page
Fax (store & forward)	Primarily one-way	~ 10 KB	Can be several minutes
Low priority transactions	Primarily one-way	< 10 KB	< 30 s

Table 18 BE Internet Application Delay Performance Objectives (ITU G.1010)

As additional studies as outlined in the previous section have been done since G.1010 was published, the delay objectives shown in Table 19 are recommended.

Game Category	Responsiveness	Consistency	Delay (ms)
FPS	Highest	Lowest	25-75
RPG	Medium	Highest	50-150
RTS	Lowest	Medium	100-500

Table 19 One way Game Delay Requirements by Category

[1] For example, for modern interactive games such as first person shooters 200 ms is not acceptable and a revised limit of 100 ms is more appropriate.

Note the delay specified is one-way and therefore aligns with research referenced above for a system response time of approximately 100 ms. Note also the system response time is measured from the perspective of the user and therefore includes application layer (game server and game client) and network layer delays.

Table 20 shows the recommendations for generic networked gaming applications in the format of G.1010. Note the delay specified is one-way and therefore aligns with research referenced above for a system response time of approximately 100 ms. Note also the system response time is measured from the perspective of the user and therefore includes application layer (game server and game client) and network layer delays.

Application	**Degree of symmetry**	**Typical amount of data**	**One-way delay or response time**	**Jitter**
Interactive games	Two-way	< 1 KB	< 50 ms	<10ms

Table 20 Interactive Gaming Delay Objectives

Some guidelines for transport parameters can be deduced from the figures above.

The minimum transmission rate of the DSL section can be computed from the ratio of the amount of data and the maximum delay. So, for example for acceptable bulk data transfer the DSL link must at least have a rate of about 8x10MB/60s = 1.333Mbit/s and preferably 5.333Mbit/s.

As mentioned earlier the use of TCP means that the throughput rate will be governed by the window size and the round trip delay. It is not uncommon to have a round trip delay to a site on the Internet of the order of 100 ms. The maximum window size in TCP is 65 KB[2] (although it is usually set lower) and so the maximum throughput obtainable in that case is just over 5 Mbit/s. This suggests that providing very high rate access transmission rates may not actually be of much use for best-effort Internet applications unless the DSL is supporting multiple simultaneous download sessions or the window scaling option is used. There are, however, potential problems with too large a window size.

9. Summary of Transport Layer QoE Recommendations

A summary of the transport layer QoE dimensions and key recommendations for each service are provided below.

[2] Unless window scale option (RFC1323) is used.

9.1 Summary Video QoE Recommendations

The figures below provide a graphical representation of the transport layer packet loss performance for video services.

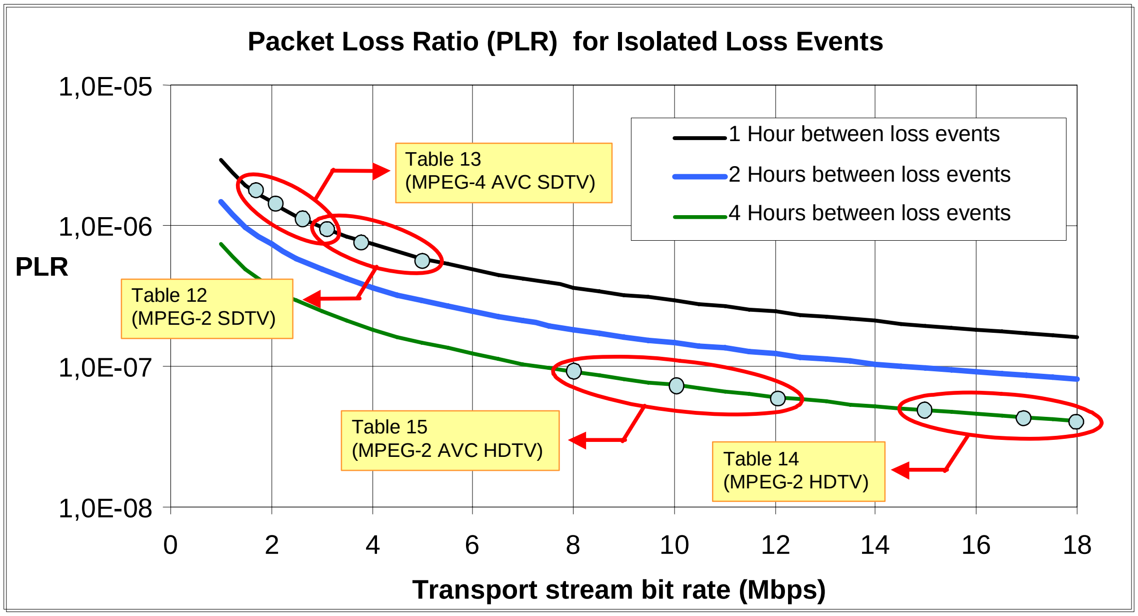


Figure 24: PLR required to meet average time between loss events of 1, 2, and 4 hours assuming isolated lost packets.

9.2 Summary Voice QoE Recommendations

There are four key VoIP QoE impacting factors:

- delay (including delay variation or jitter),
- the speech codec,
- cell/packet loss,
- echo.

Based on the ITU E model and R transmission quality rating, it has been determined that a difference of 3***R*** is not noticeable by typical users and, therefore, a triple-play service offering should be engineered within this margin in order to provide an equivalent replacement technology (ex. PSTN). Differences of 3-7***R*** might be noticeable, but most likely acceptable. Larger ***R*** degradations (greater than 7***R***) are more likely to be noticeable and should be avoided. The table below lists the recommended end-to-end QoE voice service performance objectives in triple-play service deployments.

<table>
<tr><th rowspan="2">Services</th><th colspan="2">User QoE Performance Targets</th></tr>
<tr><th>QoE metrics</th><th>QoE targets</th></tr>
<tr><td rowspan="8">Conversational voice
(CBR and VBR)</td><td colspan="2">Conversational Voice</td></tr>
<tr><td>R-factor</td><td>ΔR PSTN - Packet < 3R</td></tr>
<tr><td>delay</td><td>< 150 ms</td></tr>
<tr><td>distortion</td><td>Ie < 3R</td></tr>
<tr><td colspan="2"></td></tr>
<tr><td colspan="2">Path Interruptions Due to Failure</td></tr>
<tr><td>Frequent Interruption</td><td>80ms (affects speech intelligibility)</td></tr>
<tr><td>Infrequent interruption</td><td>3 sec (perceived as call drop)</td></tr>
</table>

Table 21 Summary of Voice QoE Guidelines

Objectives are e2e and recommended as a minimum to ensure satisfactory service quality (QoE) and therefore successful services.

9.3 Summary Best-effort Application QoE Recommendations

This document examined the web-browsing and interactive gaming applications that are most common representatives of best-effort data services offered. The table below provides the performance guidelines for these services.

BE Application	Degree of symmetry	Typical amount of data	Round trip system response time	Jitter
Interactive games	Two-way	< 1 KB	< 100 ms	<10ms
Web browsing	Primarily one-way	< 100 KB, typically 10KB	Preferred < 2 seconds / page Acceptable <4 seconds	n/a

Table 22 Summary of Best-effort Application QoE Guidelines

Appendix I – Subjective Experiments in Video Quality

I-1 Results of Subjective Experiments on Video Quality

Subjective studies have been done to determine the effects of various impairments on user perception of video quality. A study commissioned by a European IPTV service was designed to evaluate the effects of IP packet loss on user quality of experience for a standard definition (SD) IPTV service.

As shown below, the user experience of video viewing is more affected by the frequency of errors than the length of the single error itself. The study commissioned, demonstrated that controlling for the number of packets lost, viewers rated shorter, more frequent losses (low period, low distance) to be more objectionable than longer, less frequent losses (longer period, longer distance), provided the length of the error event period is reasonably short (a few tens of milliseconds).

Overall results summarized below show the influences identified in this study by all the components making up the video path, for example, an STB error concealment mechanism.

- Results show that the viewer rating is affected by (1) the frequency of error events and (2) the duration of error events.
- When the number of lost packets [amount of lost data] is held constant, viewer rating was more strongly affected by the frequency of error events than by the duration of such events (for events up to in the range of 10’s of ms in duration).
- Considering current network technologies and deployment for IPTV, IP packet loss is by far the main network transport parameter that affecting the video quality and experience along the delivery to the customer.
- IP packet loss can serve as an indicator of the quality of the video delivered to the end user (assuming the input video quality (application layer performance) to the IP network is sufficient) as well as of the general network behavior.

Metrics for packet loss behavior in the network should include both a measure of packet loss rate and a measure of mean time between loss events and loss duration to properly assess the service delivery.

I-2 Experiment details

As far as IPTV SD video service type is concerned, significantly challenging operation conditions were tested: 4 Mbps MPEG2 encoded sport (football / soccer) events.

Expert viewing sessions were run before subjective assessments and it was measured that nearly all single packet losses produced a visible error: out of 169 single IP packet losses inserted in video test streams 154 did cause a visible error; roughly in the test environment there was a 91% probability that a single packet loss produced a visible artifacts

I-2.1 Subjective test sessions specifications

Subjective tests were run partially following ITU-R BT 500-11 methodology. Test specifications are summarized here below:

- test sequences length: one minute
- a 0-10 quality rating scale was used
- a total of 6 different one minute test sequences were selected to be used in the subjective testing.
- each test sequence was tested at least against 10 different loss profiles.
- each loss profile was inserted in at least three different test sequences
- a summary of loss profiles included is:
    - Number single loss events inserted in the one minute test sequences: 1, 2, 4, 6, 10, 15, 25
    - Burst loss profiles for consecutive losses:

Loss Period	Loss Frequency/Repetition
2	1,2
3	1,3
5	1,2,3
8	1

    - Other packet loss profile tested (X = Lost R = Received):

Loss Profile	Loss Frequency/Repetition
XRXR	1
XRXRXR	1
XRXRXRXR	1
XX [Run of 24 or 36 R] XX	1

- Encapsulation: TS MPEG over UDP, 7 188 Bytes MPEG packets carried in one IP packet
- GOP = 12

I-3 Test results

Following pictures summarizes the test results. The first graph shows results for subjective tests where single, isolated packet loss events were inserted in the test sequences. A marked dependency of the perceived quality from the number of impairments perceived is observable from the curve

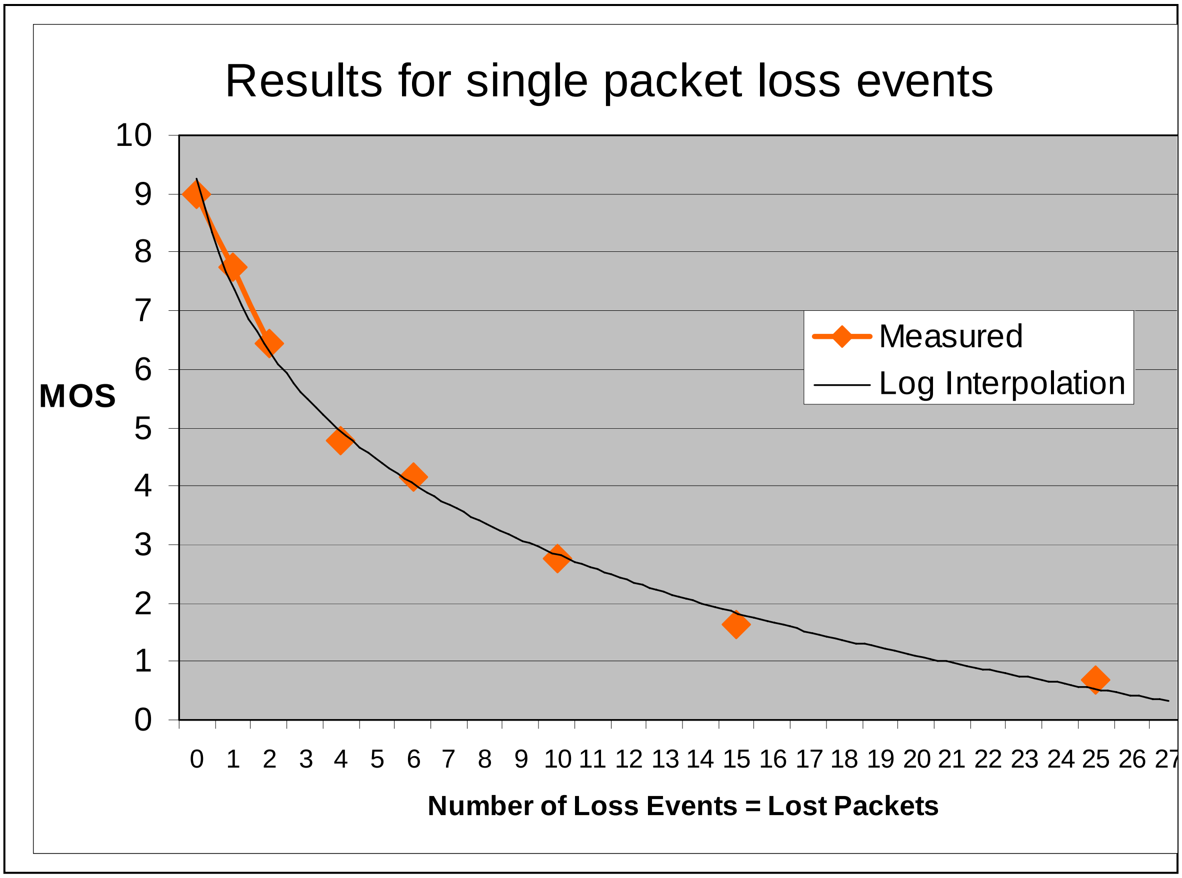


Figure I-1 Subjective test results for single packet loss events

Next figure compares subjective tests with the same number of loss events and different loss periods

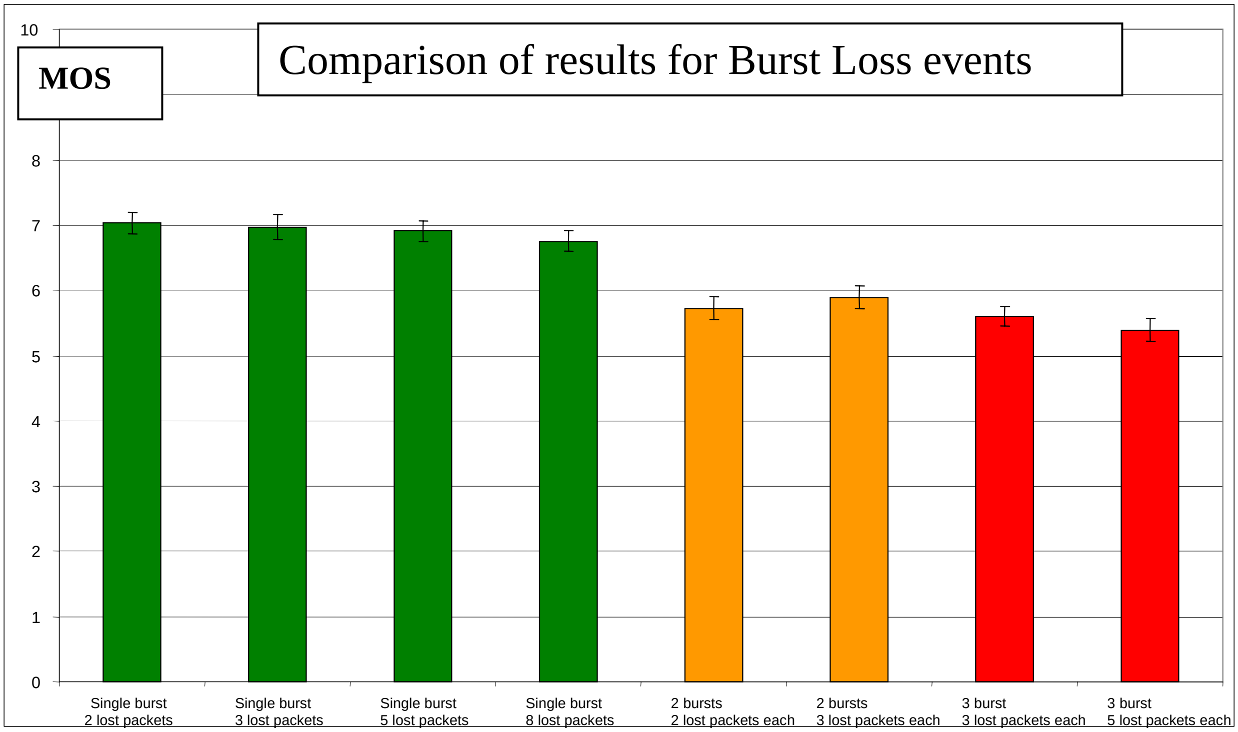


Figure I-2 Subjective test results for burst loss events

Figure I-2 highlights that even if the loss period affects the perceived quality, and longer loss periods are generally perceived as worse, the overall user experience is much more affected by the number of distinct loss events. Note that in the graph green bars represent test with single loss event, orange bars test with two loss events and red bars with three loss events. Finally, Figure I-3 shows all results together. Left vertical axis represents MOS values, while Right vertical axis represents the number of lost packets in the test.

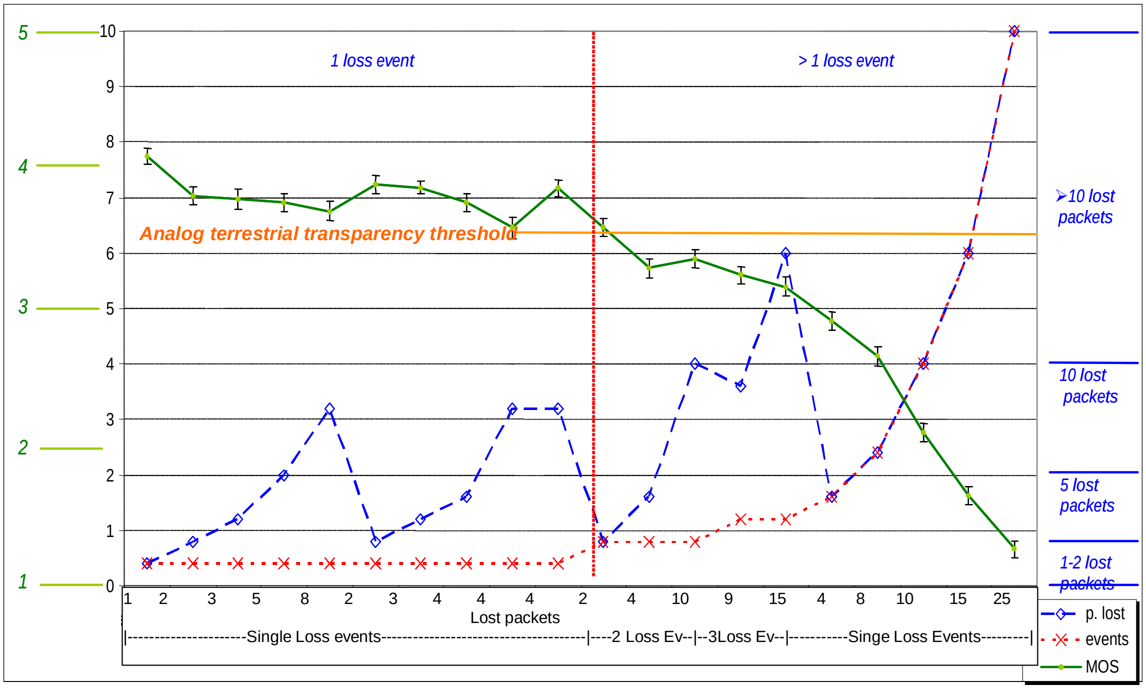


Figure I-3 Summary of video subjective results

In the horizontal axis tests results are grouped by number of loss events introduced in the test and then by overall lost packets introduced in the test. In this way, the first point of the plotted of the curves are about the results of the test where a single loss event made by a single packet loss was inserted, the second point is where a single error event was made by two consecutive lost packets was inserted, and so on, continuing then with tests where 2,3 and more loss events where introduced.

The goal of the picture is to show the correlation between quality and number of loss events and number lost packets.

I-4 Study comments

- Results shows that the user experience is more negatively affected by the frequency of error events rather than from the duration of each error event. This, at least looks valid for the maximum error length tested that was around 10s of msec.
- Considering current network technologies and deployment for IPTV, IP packet loss is by far the main network transport parameter that impacts the video quality and experience along the delivery to the customer.
- IP packet loss is then both a measure both of the video quality delivered to the end user, (this means that input video quality into the IP network is known and meets requirement stated in the application layer requirements outlined above) and of the general network performance.

As far as video service delivery is concerned, beyond absolute packet loss ratio measured, the mean time between errors events and error event duration should be considered in order to properly assess the service delivery.
Same absolute average packet loss ratio might result in quite different MOS for the service if the total amount of loss was introduced in a single error event or it was spread in different shorter error events. As shown, the latter condition is much more perceived as negative by the user.

Note that tests didn't focus on severe error events, i.e. error events whose length exceeds 200msec and goes up to seconds.

A maximum frequency for severe error events should be set in any case so that IPTV user still perceives the SD IPTV service as on par with services offered by alternate mechanisms such as cable or satellite.

Appendix II - Error Protection Mechanisms Overview

While there are many impacts to video service QoE, perhaps the most critical and difficult to control is the impact of packet loss, particularly in the access and home network segments. This section discusses the problem in greater detail and provides informational guidelines on mechanisms that may be used to mitigate the QoE impacts of packet loss on video services.

II-1 Factors influencing BB access solution performance

This section highlights some key considerations for optimizing broadband access solution performance with recommendations to be established in other DSL Forum work groups.

- Access loop BER performance and mitigation techniques
  - Base BER performance designed in to the various xDSL loop standards
  - Mechanisms to improve BER performance such as FEC and interleaving
- Level of subscriber aggregation
  - As the number of aggregation level is increased, QoS flows treatment might be impacted at the aggregation points hence disrupting the QoS granularity (per subscriber, per flows…) treatment effectiveness. If such QoS granularity is desired, extra levels of scheduling hierarchy must be implemented – see below
- Service quality requirements – soft vs. hard QoE, fairness
  - "Hard QoE": Deterministic, no-oversubscriptions of guaranteed traffic
Simpler to provision and engineer although might not be utilizing resources efficiently
  - "Soft QoE": Based upon statistical end user traffic profiles
Make assumptions on expected traffic probabilities
More complex but could utilize resources more efficiently
- Scheduling discipline
  - In order to offer performance guarantees in term of delay. Jitter, throughout, packet loss etc.., a number of scheduling disciplines are required to offer different traffic class treatment. For example, real time sensitive traffic such as voice and video would typically require a strict priority scheduler to ensure low latency, although it can also be adequately served using WFQ if engineered accordingly.
- Hierarchical level
  - The use of hierarchical scheduler provides isolation between aggregation levels hence offers more robust system performance under extreme traffic loading conditions and/or skewed inter DSLAM traffic demands. It also allows to better control the level of QoS control granularity (per application, per user, per subscriber, DSLAM…)

    - Impact and requirements depend on the scheduling and rate limiting capabilities of BRAS, aggregation node and DSLAM as well as the number of aggregation level
- Congestion control & rate limit capabilities:
    - The motivation for implementing congestion control & rate limit capabilities is to control congestion, over-subscription and fairness issues. Could prevent customers attempting to surcharge the network.
- Centralized vs. distributed QoS control
    - An architecture where QoS control is centralized would typically have full visibility of all traffic injected in the access loop hence could potentially offer hard guarantees in term of service quality or QoE. In a distributed system whereby a portion of traffic is passing through the BRAS the other injected at local points could suffer in delivering hard QoE guarantees. In a distributed architecture, the rate limiting/shaping effectiveness is limited as it doesn't have full visibility of all traffic aggregated and congestion could occurs at some local aggregation points.
- BRAS traffic visibility (video by pass)
    - Video traffic could stall lower or equal priority traffic such as multimedia traffic (voice/data) if the # of simultaneous sessions is not rate limited – and vice versa multimedia traffic could stall lower priority video traffic…
- Admission control/policy control methods
    - The main purpose is to supplement QoS scheduling performed by service nodes and DSLAM to prevent high priority traffic over-subscription
    - Should be controlled & monitored at subscriber level as well as aggregation points.
- Traffic engineering & resources allocation
    - This has to be optimized based upon all the above factors and considerations

II-2 Reliability approaches

Following sections summarize reliability approaches commonly found and employed to reduce or recover the loss of information in telecommunication networks. Recent developments and ongoing works, specific for IPTV delivery, are also described

II-2.1 Classification of approaches

Approaches to the problem of reliability in multimedia services can be broadly categorized according to the layer of the OSI stack at which they are provided as follows:

- Physical, link and network layer solutions
    - In these solutions, the underlying reliability of the network is enhanced to ensure that packet loss rates seen by the application/transport layers are sufficiently low to delivery the target application reliability. Techniques in this category include network QoS management (e.g. Intserv, Diffserv, etc.), link layer retransmission (e.g. 802.11 retransmission) and physical

layer techniques (e.g. DSL physical layer forward error correction and interleaving)

- Application/Transport layer solutions
  - In this category we include all end-to-end solutions which operate above the IP layer. Such solutions may be independent of the application, such as transport layer retransmission (e.g. TCP), or linked to the application either loosely (for example application layer forward error correction) or strongly (for example redundant video encoding techniques and loss concealment).

At some level, all solutions trade bandwidth for reliability and for a given application in any given system a service provider needs to consider which of the above techniques – or more likely what mix of techniques – provides acceptable reliability at minimized cost without jeopardizing the QoE targets.

With Application/Transport layer techniques, the resources can be dedicated to specific applications, whereas physical and link layer techniques generally apply equally to all applications using the link.

II-3 Sources of packet loss

Errors can occur at any of the layers discussed above. Errors at any level may be corrected by mechanisms in any of the layers above. The allocation of the available resources for error correction (i.e. bandwidth, and time resources, which means latency) between the different layers of the system is therefore an important question for optimal operation of the system.

Physical layer errors in electrical or RF systems are generally caused by electrical interference, for example:

- in the operator’s network
- in the home network
- in a wireless home network

Such interference may be particular severe for CE devices on commercial power where transients, surges and brownouts of commercial power and the power switching events within the home can be expected to be common. Particularly the effect of electrical impulse noise on DSL systems is well known and in-home wireless networks may be severely affected by noise/interference. Such interferences if not overcome by physical layer error correction techniques, generally appears as packet loss to the layers above.

Packet discard may also occur at the network layer due to transient buffer overloads in network equipment. The buffer occupancy is to some extent controllable by the operator given adequate information about the traffic characteristics.

Even in networks with low average utilization there can be wide variations in the instantaneous traffic load over both short and long periods of time, likely due to device failures or to traffic convergence. It can be expected that such 'transient congestion' events will be observed even in well-engineered networks.

Such congestion losses should not be confused with the 'persistent congestion' which will occur if a link is permanently oversubscribed – we can assume that a service provider network is adequately engineered for the services provided on that network and so such persistent congestion should not occur.

Finally, device failures, restarts and internal faults cause outages resulting in packet loss events at the network level.

II-3.1 Physical layer

It is important to consider the different sources of bit errors on a DSL link:

- stationary noise: this is noise which is stable in time, and is typically due to stable crosstalk from other DSL or White Gaussian background noise, e.g. –140 dBm/Hz. If no coding would be used this would result in single bit errors in most cases.
- semi-stationary noise: this is due to the on-off switching of a DSL by neighboring users having a DSL in the same copper pair binder, leading to stepwise changing cross-talk . This is typically handled by using a noise margin and therefore not considered further.
- impulsive noise: this is noise, which is not constantly present, but appears in noise bursts. This can be caused by e.g. on-hook/off-hook events. A specific form of impulse noise called repetitive electrical impulse noise (REIN) originates from e.g. dimmers and coupling onto the DSL line via a bad quality home network. If no coding would be used this would result in a long burst of bit errors.

To improve DSL reach under stationary noise conditions, DSL makes use of Trellis codes (TC) and Reed-Solomon (RS) codes. Coding corrects isolated bit/byte errors, however they have the characteristic to make error events consisting of a small burst of bit errors.

Impulsive noise destroys a sequence of consecutive bits much longer than a TC or RS is able to correct. To correct against bursts of bit errors, interleaving is used in combination with RS coding. This technique transmits the code words in an interleaved fashion, such that a destroyed sequence of consecutive bits on the DSL line is spread out over multiple code words after the de-interleaving process. The number of byte errors per RS words is reduced to a level, which can then be corrected using the RS codes. One consequence of using this approach is that it introduces an additional interleaving delay. Different delay settings have been defined in DSL standards; delays range between 0 and 63 ms, but are typically 16ms.

In practice, two operation modes are commonly used: "fast mode", which does not use interleaving and hardly introduces any delay (modem one-way delay typically 2 ms) but which is prone to impulsive noise, and "interleaved mode", which can correct practically all impulsive noises, at the expense of additional interleaving delay of typically 16 ms. Together with the signal processing delay this results in a modem one-way delay of typically 20ms.

For stationary noise, TR-067 estimates for the purpose of CPE margin verification testing that a lost packet corresponds, on average, to 15 bit errors in fast mode and to 40 bit errors in Interleaved mode. This leads to the following formulae for converting between Bit Error Rate (BER) and Mean Time Between Errors (MTBE):

- for fast mode (i.e. Trellis coding only): BER = 15 / MTBE
- for interleaved mode (i.e. Trellis and Reed-Solomon coding) BER = 40 / MTBE

Since the number of consecutive bit errors is usually (much) smaller than the number of bits needed to transport an IP packet, we further assume that one error event results in the loss of exactly one packet.

II-3.2 Link layer

For links with relatively high error rates it is usual to provide some link layer reliability mechanism. Note that, as we shall discuss in more detail below, it is not sufficient for the overall end-to-end operation of the system to provide reliability on each link separately – this is because there are always sources of loss within the network elements between links which necessitate an end-to-end reliability mechanisms of some kind. As a result, link layer reliability is generally only provided on links with relatively high error rates.

Examples are mostly wireless systems. In the case of bi-directional channels, then some form of Automatic Repeat Request is usually provided (e.g. 802.11 systems, digital cellular radio networks). In case of unidirectional channels, then link level forward error (erasure) protection may be provided (for example in the 3GPP2 cellular broadcast specifications or DVB-H). Note that 802.11 does not provide link layer reliability mechanisms in the multicast case.

In the case of DSL links, then error correction is provided at the physical layer as discussed above and there is no further link layer protection.

II-3.3 Application/Transport layer

These two are considered together since they both exist only at the end-systems and thus there is scope for interaction between the two.

IP networks do not claim to provide reliable delivery of packets end-to-end and so, depending on application or service QoE requirements, there might be a requirement for end-to-end reliability mechanisms even when physical and link layer mechanisms, such as those described above for DSL are in use. Supposing that such mechanisms will be required (and thus a ‘sunk cost’), it is natural to look for opportunities to reduce costs by avoiding unnecessary additional reliability mechanisms within the network (the “end-to-end principle”).

End-to-end packet loss rates for IP networks can vary significantly. ITU-T recommendation Y.1541 recommended end-to-end packet loss rate targets of 10^{-3} for traditional data services such as web browsing and indeed little performance improvement is seen in TCP-based services generally if the packet loss rate is less than this. Provisional classes are under evaluation in order to support video services and generally very loss intolerant application, these classes define 10^{-5} as the target loss rate

Despite the IP loss rate required, considering the several kind of access technologies available (xDSL, WiFi, WiMAX, etc) and their intrinsic different performances and behavior, an end to end mechanism is an appealing solution to achieve or cooperate in offering the QoE targets to the larger customer base as possible.

End-to-end reliability mechanisms can be categorized as follows:

- loss concealment
- automatic repeat request
- forward error correction

II-4 Loss concealment

Loss concealment is an application-specific technique, which allows an apparently reliable service to be delivered to the user even in the presence of packet loss.

MPEG-2 provides error concealment techniques based on temporal and/or spatial interpolation of missing I-frames. Motion compensation information is not included by default in MPEG-2 I-frames, but syntax extensions have been developed for this purpose.

As the structure of MPEG-4 is similar to that of MPEG-2, similar MPEG processing functions should be applicable to packets missing from an MPEG-4 video stream. MPEG-4 part 2 and MPEG-4 part 10/ AVC/H.264 both use the same I,P,B-frame distinction as MPEG2.

The disadvantage of MPEG error concealment is that this additional processing introduces additional complexity in the terminal (MPEG decoder) functions, which are high volume components of the IPTV architecture. The additional processing may also introduce some additional delay to the rendering system which may impact other metrics for the user experience e.g. channel change times. Including additional information in the stream that helps with error-concealment also has a bandwidth overhead, and the tradeoffs between using bandwidth for error-concealment versus using bandwidth for other reliability mechanisms need to be considered.

The extent to which loss concealment can achieve a reliable service is highly application-specific, dependent on the encoding and decoding techniques available and indeed subjective because it depends on the definition of “reliable service” i.e. the level of ‘unconcealed’ or ‘partially concealed’ errors that is acceptable.

II-4.1 Automatic repeat request (ARQ)

ARQ protocols detect lost packets and send a request to the transmitter to repeat the packet. The most common example of this is TCP.

In general receipt of a repeat packet requires at least one Round Trip Time from the point at which the loss is detected and the repeat request sent. If losses are at a level where loss of a repeat request or the repeated packet itself must be recovered then somewhat more than two Round Trip Times are required. This is because the loss of the repeat request or repeated packet can only be detected by expiration of a timer at the receiver. Note that if the overall packet loss rate (in both directions) is, say, 10^{-3} then the probability that two repeat requests are required is around $2*10^{-6}$, which may still be too high for IP TV services, especially HD, meaning that the system must support two repeat requests. The effectiveness of the ARQ technique is thus highly dependent on the network round trip time and so it is necessary to ensure careful network design to keep RTT under control.

A benefit of this approach is that it avoids the additional overhead due to FEC at times when no payloads are lost. Therefore this approach only requires minimal network bandwidth. At the same time, since packets are only sent when needed (and assuming the loss process is random), the required overhead bit rate fluctuates more over time. If for instance a noise burst would “overpower” the DSL interleaving process, it is perfectly possible to request retransmission of a number of successive packets (e.g. 4). This means that the download rate will increase for 4 packets worth of data.
Finally ARQ protocols will hardly scale in some environments as IP multicast (SSM, one to many) transmission when a fault is causing a severe loss (even if for a short period) to a large number of customers.

II-4.2 Forward Error Correction - Application layer FEC for media streaming

Forward Error Correction at the Application/Transport layers generally refers to packet erasure correction techniques. In these techniques an amount of data is sent which is in total greater than the stream to be communicated, with the property that the stream can be reconstructed from any sufficiently large subset of the transmitted data. The stream is thus resilient to a certain amount of loss (at most the difference between the transmitted and the original data size).

In general for steaming applications there are considerable advantages to using “systematic” FEC codes, in which the original packets of the stream (“source packets”) are sent accompanied by a certain overhead of “repair” packets. The repair packets can be used to recover source packets which have been lost between sender and receiver.

Use of FEC has an impact on latency at the decoder (the time between receipt of a packet and when it can be played out). This is because the receiver must ensure that when a packet is lost, it has enough data in its playout buffer to continue playout until that packet can be recovered. In particular, if T_{AB} is the sending time difference between packets A and B, where packet B could assist in the recovery of an earlier packet A, then playout of the stream must be delayed at the receiver by at least the maximum of all such T_{AB} values for the stream.

Note that playout of the stream may be delayed at the receiver for other reasons, for example the need to wait for an I-frame to begin playout. So, the additional latency caused by the FEC may be less than the maximum T_{AB} discussed above.

Commonly, FEC protection is applied to the stream in blocks, known as protection periods. This is illustrated in Figure II-i below (note again that for a systematic code, then the original packets are included within the stream of 'encoded packets' shown here):

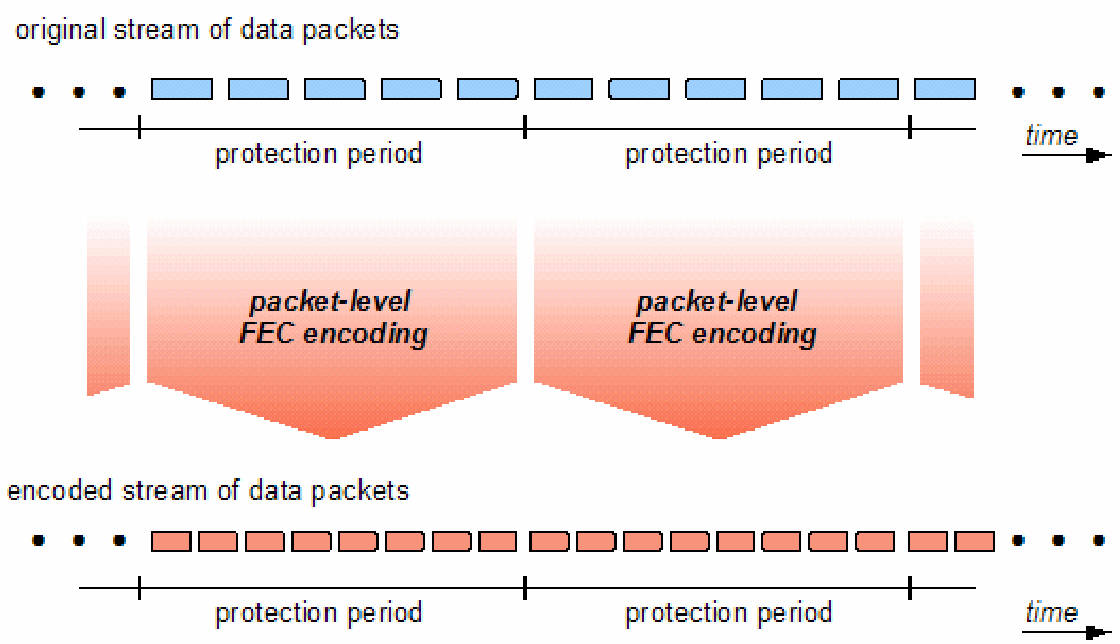


Figure II-1: FEC Protection Periods

The maximum T_{AB} is then equal to the protection period, since packets in one protection period can never assist in recovery of packets from another protection period. The protection period thus gives the minimum latency required at the receiver.

There are advantages in principle to using as long a protection period as possible. This is because the amount of FEC overhead must be engineered to be sufficient to overcome a particular 'worst case' level of losses within a protection period.

If losses are bursty, and the 'worst case' is an isolated burst of given length, then the amount of repair data required is equal to the length of this worst case burst error (or just slightly more, depending on the FEC code). So then the longer the protection period the lower the overhead required as a fraction of the source bandwidth.

If losses are random, then the 'worst case' would be given as the number of losses expected to be seen in the worst protection period within some target reliability period during which there should be no errors after FEC. In this case, the larger the protection period the closer this figure will be to the *average* losses per block and hence, again, the lower the overhead.

Figure II-ii illustrates this trade-off by showing the theoretical minimum FEC overhead required to achieve a target reliability period, or "Mean Time Between Artifacts (MTBA)", of 10^3, 10^4 and 10^5 seconds as a function of the protection period for a 3Mbit/s stream (assuming 1316 byte packet payloads) in the presence of independent random losses at a Packet Loss Rate of 10^{-3}. (The 'glitches' in this chart are quantization effects due to the assumption that the FEC overhead is a whole number of packets of the same size as the media packets). It is important to note that this chart simply illustrates the theoretical effect of increased protection period in the presence of independent random packet loss – in practice packet losses are not independent.

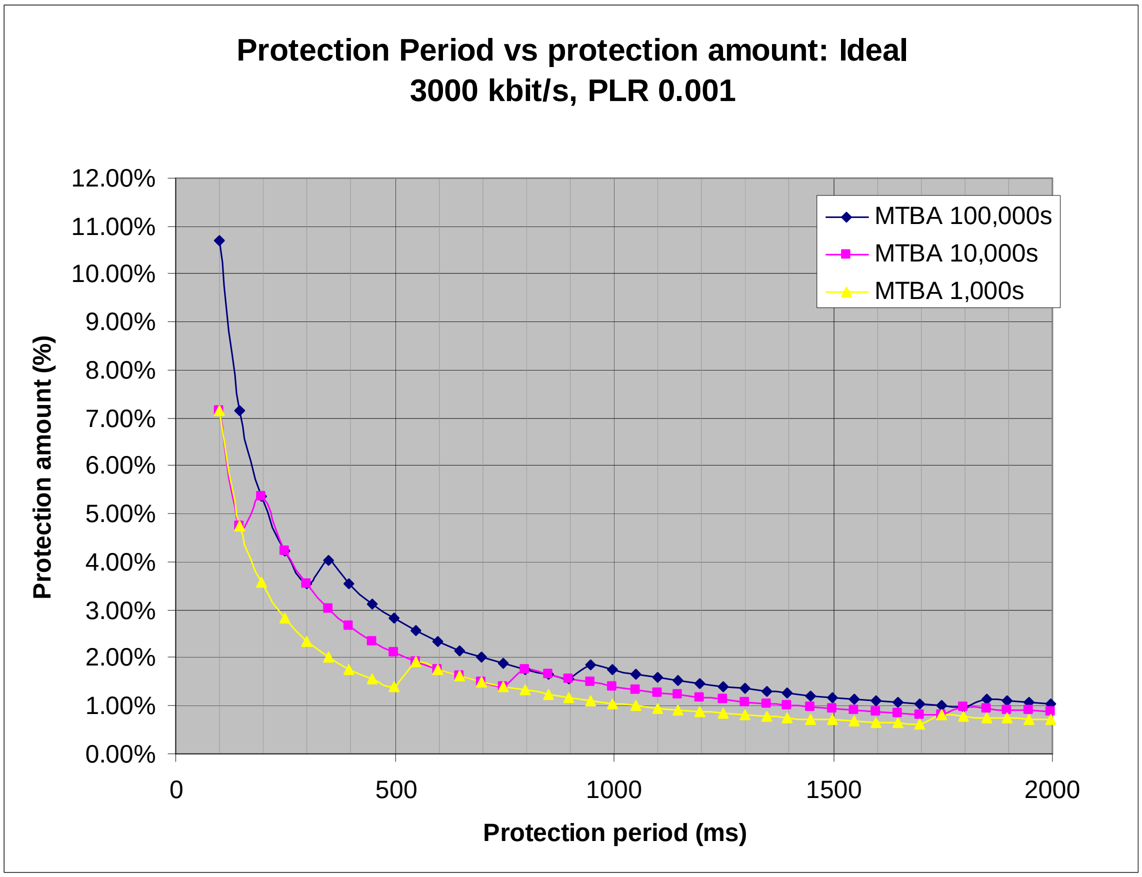


Figure II-2: Protection amount vs. Protection Period: independent random packet loss

An advantage of applying FEC at the application layer is that parameters such as the protection period can be adapted to the application. Additionally, in the case of video, there is considerable advantage if the FEC protection periods can be synchronized with the video frame structure, for example so that each protection period begins at an I-Frame. This not only provides protection on related quantities of data, but also can help to reduce significantly the additional latency introduced by FEC.

II-5 Existing standards/activities (informative)

Many possible FEC schemes for streaming media exist which could be applied to video. Currently only three packet erasure correction schemes for streaming media that have been standardized elsewhere:

- IETF RFC2733
  - This defines a simple mechanism for applying short block parity codes to RTP streams. The scheme is limited by the small number of packets that can be protected as a block (24 packets). This RFC has not been widely implemented and will likely soon be obsoleted by an update which provides slightly longer blocks (48 packets) and the possibility to apply unequal protection to different parts of each packet.
- 3GPP TS26.346 "Multimedia Broadcast/Multicast: Protocols and Codecs"
  - This standard defines a generic framework for application of FEC to media streams. The framework is not specific to RTP and operates just above the UDP layer. This framework could be used with many FEC codes, however 3GPP specify and require support of a single specific code (the Digital Fountain Raptor code).
- ETSI EN301 192 "Digital Video Broadcasting: DVB Specification for Data Broadcasting"
  - This defines what is in fact a link layer erasure code intended to be used with the DVB-H system for transmission to mobile terminals. This FEC scheme operates below the IP layer and so is completely independent of the applications and is based on a large Reed-Solomon erasure code.

The IETF has recently initiated a new working group ("fecframe") to standardize a framework for application of FEC to media streams along similar lines to that defined by 3GPP. The framework will not specify a particular FEC code but will use an approach similar to that adopted by the IETF RMT working group which standardized protocols for reliable file delivery over IP multicast. The RMT group defined an FEC Building Block which described how the specification of protocols which use FEC could be separated from specification of the FEC codes themselves. This results in a set of "plug & play" specifications which can be combined according to the needs of a given application.

Finally, the DVB IP Infrastructure group (DVB-IPI) has begun an activity to evaluate and select an FEC code for IP TV. Currently this group has agreed a set of evaluation criteria and two proposals are under evaluation: the Digital Fountain Raptor code and a short block parity code published by the Pro-MPEG Forum (generally referred to as "Code of Practice 3", or 'CoP3').

II-6 Considerations for AL-FEC selection/design (informative)

There are many factors to consider in the selection or design on a Forward Error Correction code for video applications. Some of these are:

- Error correction performance
  - Figure II-ii above shows the theoretical best possible performance for random losses. Practical FEC codes do not in general meet this theoretical ideal: some additional overhead is required to achieve a given quality target and the amount of this overhead can vary significantly between different FEC codes.

- Computational complexity and memory requirements
  - Complexity should be sufficiently low that the code can be implemented in software on resource-constrained Set-Top Boxes.

- Flexibility
  - It is desirable to be able to tailor the FEC protection provided to the application and network conditions. For example, it would be desirable to be able to choose the size of the protection period (in ms) and the amount of protection (as a % overhead) independently according to the latency which can be tolerated and the level of errors which must be overcome respectively. The performance and complexity of the code should remain as good as possible independent of these choices.

    Furthermore, in DSL deployments the level of protection required may vary from user to user. In the unicast case, it is desirable to be able to flexibly set the protection amount for each user based on real-time feedback about the experienced losses. In the multicast case, it is desirable to be able to filter out some of the “repair” packets so that they do not unnecessarily consume bandwidth on lines with low losses. In both cases then, again, the performance and complexity of the code should remain as good as possible independent of these choices.

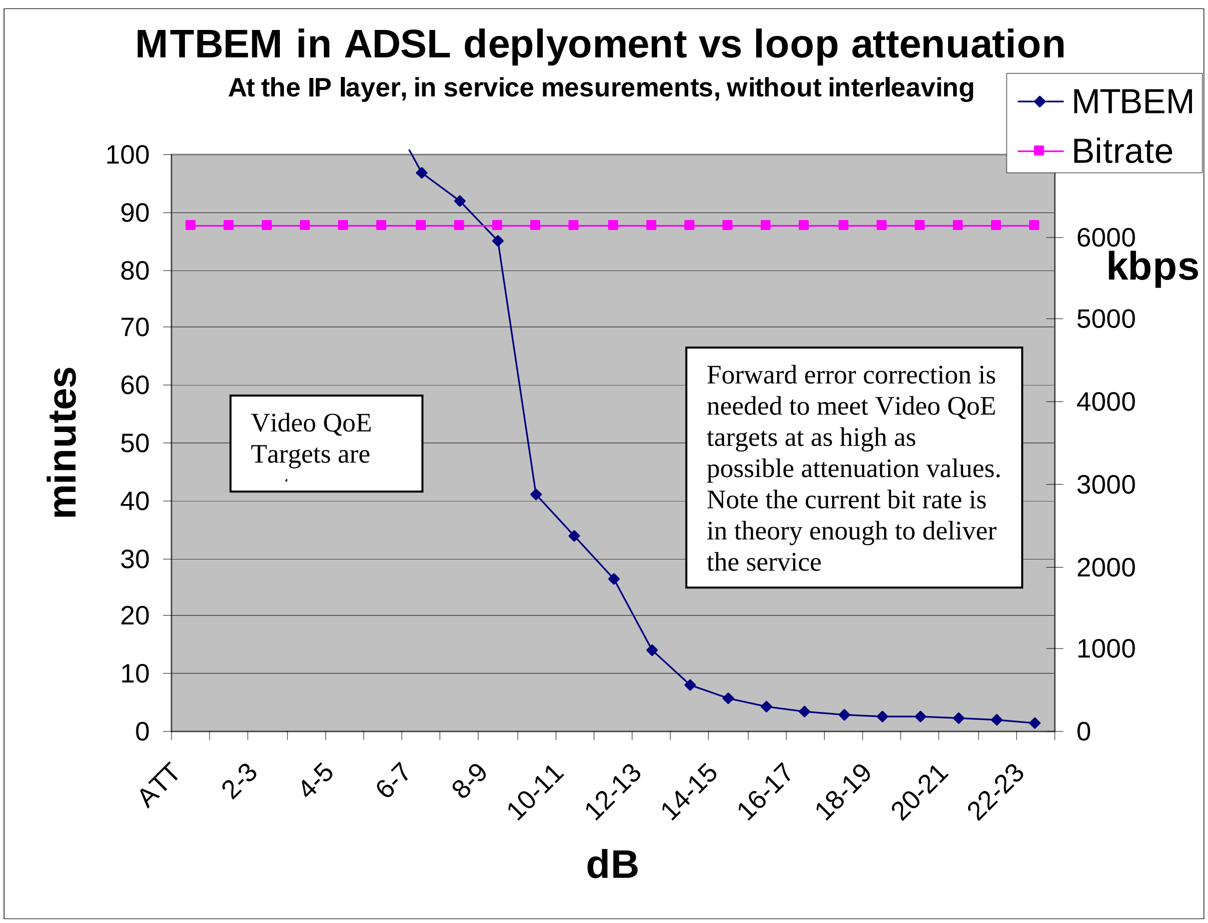
    Finally, in cases where several streams are to be delivered to the same receiver. (The simplest example would be where the audio and video components of a service are provided in separate RTP streams.) In this it is considerable more efficient to provide FEC protection across the combined data of all the delivered streams, rather than independently.

II-7 Extending the reach to deliver video services and improving the end-to-end performance to achieve QoE targets in the xDSL environment – Forward Error Correction

In this section, specific xDSL features and cases are analyzed in order to offer guidelines to achieve QoE targets considering practical copper loop behavior and the need to control its performance overtime as much as possible
DSLs technologies differ in several transmission and design aspects. Regardless of the xDSL specifications deployed, some common properties and constraints exists in the DSL

access, one of the key elements is the performance and achievable speed being dependent on the copper quality and .length
Given a certain minimum bandwidth available needed to deliver the services as engineered by the operator, a challenging task is to extend as far as possible the loop length that can sustain the bandwidth needed for the delivery of all the offered services and meet at the same time the desired QoE targets.
Generally, considering a fixed bit-rate configured in the access loop, the probability that the noise corrupts the DSL signals increases with the length (and the attenuation) of the loop and even a small number of bit errors will cause the loss of a whole IP packet transporting up to 1460 Bytes of data.
Forward Error Correction techniques or other error recovery methods can be used to reduce the impact of transmission errors, Forward Error Correction can be applied at the Physical layer or at the application layer.

The following pictures shows the need to employ such FEC mechanism to extend the QoE offered to the user in a ADSL access. The same concepts will apply for a VDSL or ADSL2+ deployment. The higher bandwidth available through these standards will be used also to augment the premium service amount of offering (mix of more than one SD and HDTV channels offered to the customers.)



Note that MTBEM is Mean Time Between Errored Minutes, where an Errored Minute is a minute were at least one error is measured. Errors are measured at the IP layer so that at least one error means at least one IP packet loss in the video stream being played out by the STB.
No
FEC was used in the chain and error protection is the only one provided by the base DSL settings (FAST Path)
The graph shows that as the loop length increases the service quality decreases even if the bandwidth available is the same for all the loops considered. This is likely due to the higher sensibility to the noise of long loops were transmission signal is attenuated at the receiver side. This kind of noise is likely to be the one present in the loop due to interference from other signals present in the cable.

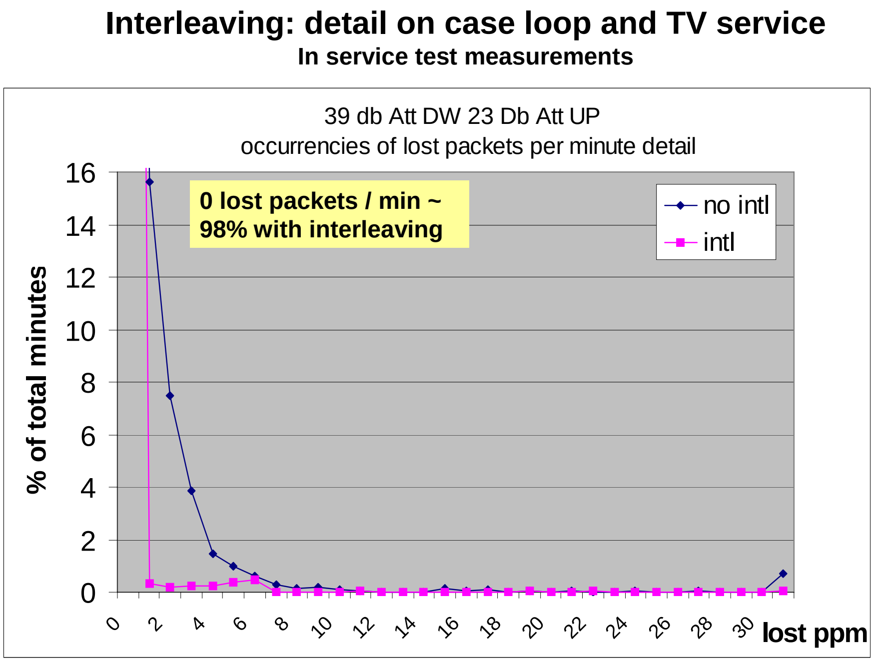
II-7.1 ADSL – PHY Layer FEC improvements through interleaving for attenuated loops which support bit rate for Video Services

Here it is shown an example of a loop behavior with and without interleaving as specified by the ADSL standard. Interleaving was set to 16msec. Case is chosen to represent a long loop as far as video service delivery over ADSL is concerned and measurements were done in real network environment.
Next table shows the number of lost packets per minutes before and after the application of the interleaved path. Noticeably nearly all the (frequent) kind of errors (lost packets) which were measured with the FAST Path settings were recovered with use of interleaving and RS settings as specified in ADSL standard for a 16msec latency path.
Measured loop upstream attenuation is 23dB, lost packets are measured in service through STB reporting

Interleaving:
Case of field loop. packet loss compared before and after the activation of interleaving
23 Db Attenuation UP occurrences lost packets per minute

NO Intereleaving				Interleaving			
Lost PPM	***Frequency***	***Lost Pckts***	***%***	***Lost PPM-After***	***Frequency***	***LostPckts After***	***%***
0	1474	0	68.08314	0	3210	0	98.01527
1	338	338	15.61201	1	11	11	0.335878
2	162	324	7.482679	2	6	12	0.183206
3	84	252	3.879908	3	7	21	0.21374
4	32	128	1.47806	4	8	32	0.244275
5	21	105	0.969977	5	13	65	0.396947
6	13	78	0.600462	6	16	96	0.48855
7	6	42	0.277136	7	0	0	0
8	3	24	0.138568	8	0	0	0
9	4	36	0.184758	9	0	0	0
10	2	20	0.092379	10	0	0	0
11	1	11	0.046189	11	1	11	0.030534
12	0	0	0	12	0	0	0
13	0	0	0	13	0	0	0
14	0	0	0	14	0	0	0
15	3	45	0.138568	15	0	0	0
16	1	16	0.046189	16	0	0	0
17	2	34	0.092379	17	0	0	0
18	0	0	0	18	0	0	0
19	1	19	0.046189	19	1	19	0.030534
20	0	0	0	20	0	0	0
21	1	21	0.046189	21	0	0	0
22	0	0	0	22	1	22	0.030534
23	0	0	0	23	0	0	0
24	1	24	0.046189	24	0	0	0
25	0	0	0	25	0	0	0
26	0	0	0	26	0	0	0
27	1	27	0.046189	27	0	0	0
28	0	0	0	28	0	0	0
29	0	0	0	29	0	0	0

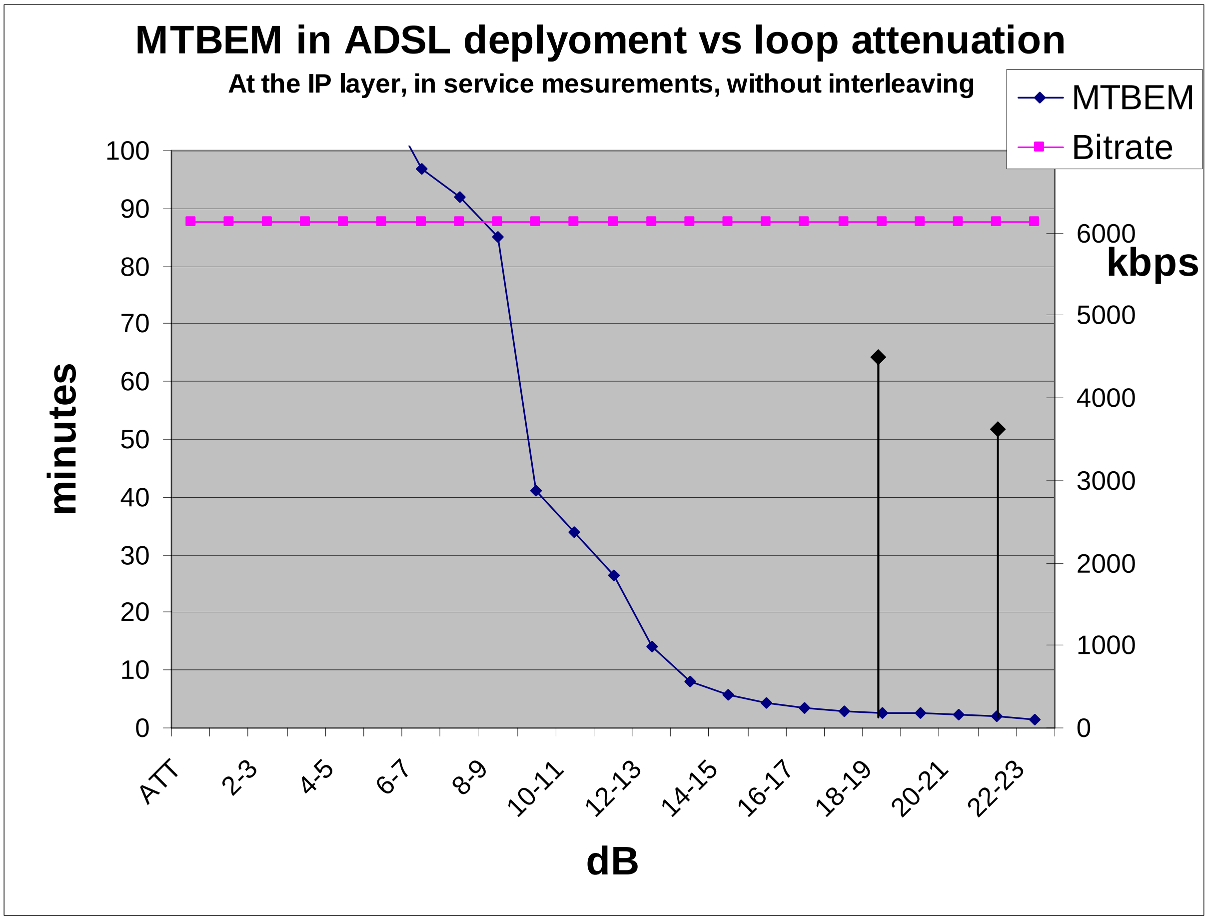
Previous table is graphed in the next picture where the behavior difference of the loop in the two configurations is easily captured.



Note the peak observed around 3-6 packets per minute with the interleaved path. This likely due the non linear behavior of the error correction scheme when the impulse noise interval exceeds correction capabilities. For those minutes it is likely that the 6 lost packets were distributed in error events with more than one lost packet.

It is not the scope of this analysis to state that this case loop meets QoE targets, rather to point-out the gain that is achieved with the ADSL 16 milliseconds interleaved path. Following values summarize the improvement done on error resilience for this case:

PLR before: 2X10-3
PLR after 2.5X10-4
MTBEM before: 3 minutes
MTBEM after: 50 minutes
Considering other test and measurements done on these kind of loops it is next shown the picture XX with the gain obtained through ADSL interleaved path



Interleaving path usage clearly helps in meeting QoE targets, therefore the use of 16 ms interleaved path is suggested when offering triple play services with SD IPTV over ADSL. 16mseconds of delay introduced with the interleaved path might impact interactive services with tight delay requirements such as gaming. In this case interleaving approach might not be applicable

Appendix III – Gross Error Detection

Gross errors in this context are defined as video impairments that cause either freeze frames (video frame is repeated more than once) or dropped frames (video frame was not played out). Preliminary user testing was done in the context of a wireless LAN application to assess user's rating of these types of gross errors. The results are presented here for informational purposes.

III-1 User Assessment of Gross Error QoE Impacts

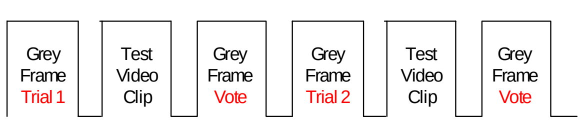
- Video sequences with varying levels of degradation were generated to provide a representative sample of video streaming playback errors
- A series of 240-frame Standard Definition (720x486) clips was assembled into a one-hour test session (10 video clips x 5 conditions x 2 Error types)
- A panel of non-expert subjects (50% male and 50% female) were asked to evaluate each clip for "playback smoothness and fluidity," not their opinion of the video content
- Before each test session, subjects were shown a sample of the best and worst clips to establish a frame of reference in order to reduce the impact of participant inherent biases

Subjective Video Quality Scale

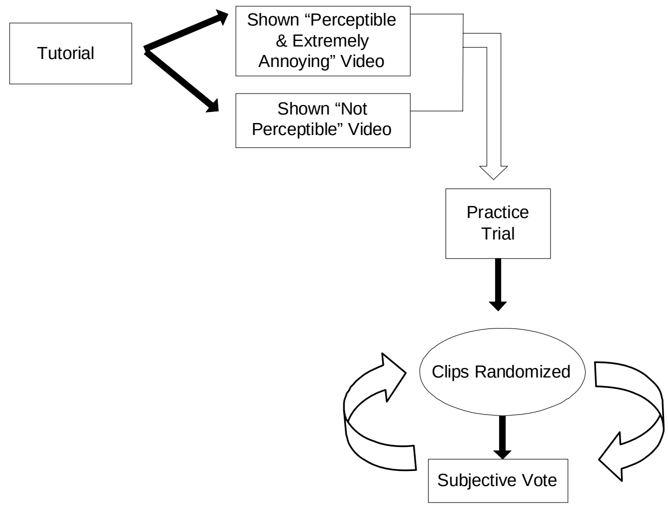
After each video sequence, participants were presented with a choice of five adjectives describing their opinion of the video experience on a subjective scale

Numerical Value	User Experience
5	Not Perceptible
4	Perceptible but Not Annoying
3	Perceptible & Slightly Annoying
2	Perceptible & Very Annoying
1	Perceptible & Extremely Annoying

Trial Structure

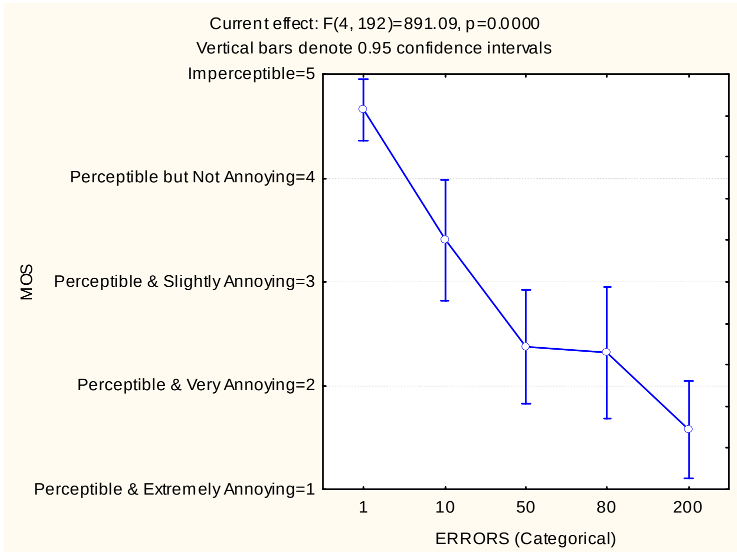


User Experience Assessment – Workflow



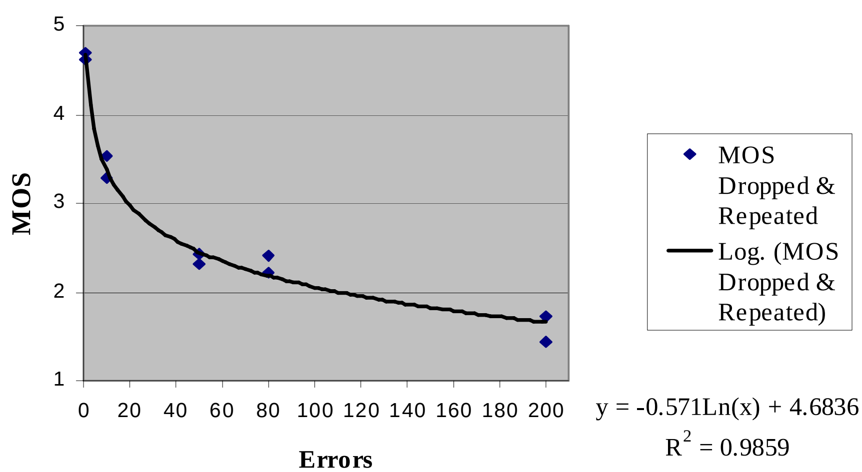
II-1.1 User Assessment of Gross Errors Results

Collapsed Across Dropped and Repeated for Each Condition



Combined Mean Opinion Score for Dropped and Repeated Frames

MOS Log fit for both Dropped and Repeated Errors



Log fit graph for predicting dropped and/or repeated errors with high correlation.

III-2 QoE Guidelines for Gross Video Errors

Dropped Frames			Repeated Frames			Dropped & Repeated Frames		
Experience	Errors	Percent	Experience	Errors	Percent	Experience	Errors	Percent
5	0.7	0.3%	5	0.5	0.2%	5	0.6	0.3%
4	3.8	1.6%	4	2.9	1.2%	4	3.3	1.3%
3	20.4	8.5%	3	17.7	7.4%	3	18.9	7.8%
2	110.3	46.0%	2	109.5	46.0%	2	109.9	46.0%
1	595.4		1	677.4		1	633.4	
Based on 8 sec clip = 240 frames								

IV Bibliography

[1] TR-058 Multi-Service Architecture & Framework Requirement, DSL Forum TR-058, Sept 2003
[2] TR-059 DSL Evolution - Architecture Requirements for the Support of QoS-Enabled IP Services
[3] "Base Requirements for an ADSL Modem with Routing", DSL Forum TR-068,May 2004
[4] "CPE WAN Management Protocol", DSL Forum TR-069, May 2004
[5] "Multi-Service Delivery Framework for Home Networks", DSL Forum TR-094,
[6] "Migration to Ethernet Based DSL Aggregation", DSL Forum, TR-101v6, Feb. 2005.
[7] "Service Interface Requirements for TR-058 Architectures", DSL Forum, TR-102v3, Aug. 2004.
[8] "Communications quality of service: A framework and definitions", ITU-T SERIES G: TRANSMISSION SYSTEMS AND MEDIA, DIGITAL SYSTEMS AND NETWORKS Quality of service and performance, Recommendation G.1000, November 2001.
[9] "End-user multimedia QoS categories", ITU-T SERIES G: TRANSMISSION SYSTEMS AND MEDIA, DIGITAL SYSTEMS AND NETWORKS Quality of service and performance, Recommendation G.1010, November 2001.
[10] "Performance parameter definitions for quality of speech and other voiceband applications utilizing IP networks". ITU-T SERIES G: TRANSMISSION SYSTEMS AND MEDIA, DIGITAL SYSTEMS AND NETWORKS, Quality of service and performance - Generic and user related aspects, Recommendation G.1020, November 2003.
[11] "The E-model, a computational model for use in transmission planning.", ITU-T SERIES G: TRANSMISSION SYSTEMS AND MEDIA, DIGITAL SYSTEMS AND NETWORKS, International telephone connections and circuits – General definitions. Recommendation G.107, March 2003.
[12] "Network performance objectives for IP-based services", ITU-T SERIES Y: GLOBAL INFORMATION INFRASTRUCTURE AND INTERNET PROTOCOL ASPECTS, Internet protocol aspects – Quality of service and network performance, Recommendation Y.1541, May 2002 (and related updates - 2006)
[13] "Quality of Service ranking and measurement methods for digital video services delivered over broadband IP Networks", ITU-T Study Group 9, J.241 Draft, February 2005: http://www.itu.int/itudoc/itu-t/aap/sg9aap/recaap/j241/
[14] J.140 Subjective picture quality assessment for digital cable television systems,
J.141 Performance indicators for data services delivered over digital cable television systems,
J.142 Methods for the measurement of parameters in the transmission of digital cable television signals,
J.143 User requirements for objective perceptual video quality measurements in digital cable television,
J.144 Objective perceptual video quality measurement techniques for digital cable television in the presence of a full reference,

J.145 Measurement and control of the quality of service for sound transmission over contribution and distribution networks,
J.146 Loop latency issues in contribution circuits for conversational TV programs
J.147 Objective picture quality measurement method by use of in-service test signals
J.148 Requirements for an objective perceptual multimedia quality model
J.149 Method for specifying accuracy and cross-calibration of Video Quality Metrics (VQM)

[15] Video Quality Experts Group (VQEG), FRTV, RRNR-TV and HDTV working groups: www.vqeg.org , http://www.its.bldrdoc.gov/vqeg/projects/projects.php

[16] Professional MPEG Forum, http://62.73.167.57/index.html

[17] Digital Living Network Alliance – home and technical working groups: http://www.dlna.org/home ; http://www.dlna.org/about/organization/technical/

[18] Video Services Forum (VSF), Technical AdHoc groups: http://www.videoservicesforum.org/resources/index.htm

[19] "CableLabs® Video-On-Demand Content Encoding - Profiles Specification", CableLabs® document MD-SP-VOD-CEP-I01-040107, January 2004: http://www.cablelabs.com/projects/metadata/downloads/specs/MD-SP-VOD-CEP-I01-040107.pdf

[20] IP Performance Metrics (IPPM) Charter, Internet Engineering Task Force (IETF) http://www.ietf.org/html.charters/ippm-charter.html , including RFC3357 One-way Loss Pattern Sample Metrics: http://www.ietf.org/rfc/rfc3357.txt

[21] " Network Performance and QoS Subcommittee (NPQSC) (formerly T1A1.3)", Alliance for Telecommunications Industry Solutions (ATIS) workgroup: http://www.atis.org/0010/npqsc.asp

[22] IPTV Interoperability Forum (IIF), Alliance for Telecommunications Industry Solutions (ATIS) workgroup: http://www.atis.org/iif/index.asp

[23]"Speech Processing, Transmission and Quality Aspects (ETSI STQ)", European Telecommunications Standards Institute (ETSI), http://portal.etsi.org/portal_common/home.asp?tbkey1=STQ

[24] "Telecoms & Internet converged Services & Protocols for Advanced Networks – TISPAN", European Telecommunications Standards Institute (ETSI), http://portal.etsi.org/Portal_common/bottom.asp?Register=&tbid=625&SubTB=625%2C+626%2C+627%2C+628%2C+629%2C+630%2C+631%2C+632%2C+633&TAB_ID=&Param=

[25] "DVB-IP Project", Digital Video Broadcasting Project, http://www.dvb.org/technology/white_papers/wp12.DVB-IPTV.pdf

[26] ITU-T Study Group 12 - Performance and quality of service. Question 12 Study Period 2005-2008: http://www.itu.int/ITU-T/studygroups/com12/index.asp

[27] ITU-T Study Group 12,Question 13/12 - Multimedia QoE/QoS performance requirements and assessment methods: http://www.itu.int/ITU-T/studygroups/com12/sg12-q13.html

[28] RFC-3984 - RTP Payload Format for H.264 Video, Internet Engineering Task Force (IETF) http://www.ietf.org/rfc/rfc3984.txt?number=3984.

---

[29] ANSI / TIA-921 "Network Model for Evaluating Multimedia Transmission Performance over Internet Protocol", 2006.
[30] H. de Ridder et al.(1995). Naturalness and image quality: Chroma and hue variation in color images of natural scenes, in Proc. SPIE , vol. 2411, 51--61, San Jose, CA.
[31] Winkler, Stefan (1999). Issues in Vision Modeling for Perceptual Video Quality Assessment, Signal Processing 78(2):231–252.
[32] ITU-R BT.500-11. Methodology for the Subjective Assessment of the Quality of Television Picture, 2002. ITU-R Recommendations
[33] " Enhanced IP-Based Video QoS Performance Objectives", Network Performance and QoS Subcommittee (NPQSC) (formerly T1A1.3), Alliance for Telecommunications Industry Solutions (ATIS) workgroup:
http://contributions.atis.org/upload/PRQC/PRQC/3A100930.doc
[34] Nielsen, Jakob (1994). Response Times: The Three Important Limits. Usability Engineering, Chap. 5, AP Professional, Boston, MA, 1994. Also online at:
http://www.useit.com/papers/responsetime.html

[35] Bickford, Peter (1998). Human Interface Online: Worth the Wait?, Netscape Communications, Netscape Viewsource Magazine.

[36] MIL-STD-1472C, in Brown, C. Marlin (1988). Human Computer Interface Design Guidelines, Abex Publishing, Norwood NJ, 115-119.

[37] Bellcore specs on availability: TA-NWT-000418 and TA-NWT-000909
[38] Telcordia Technologies, formerly known as Bellcore, was created during the divestiture of the Bell System in 1984 to serve the Bell operating companies. www.telcordia.com.
Relevant reliability documents include:

- Telcordia (Bellcore) GR-512-CORE, Issue 1, January 1995, LSSGR Section 12: Reliability www.telcordia.com (Informative Reference)
- Telcordia (Bellcore) GR-499-CORE, Issue 2, December 1998, Transport Systems Generic Requirements (TSGR): Common Requirements (Informative Reference)
- Telcordia (Bellcore) TR-TSY-000511, Issue 2, July 1987, LSSGR Section 11: Service Standards (Informative Reference)
- Telcordia (Bellcore) TA-NWT-000909, Issue 2, December 1993, Generic Requirements and Objectives for Fiber in the Loop (FITL) Systems (Informative Reference)
- Telcordia (Bellcore) TR-NWT-000057, Issue 2, January 1993, Functional Criteria for Digital Loop Carrier (DLC) Systems (Informative Reference)
- Telcordia (Bellcore) GR-30-CORE, Issue 2, December 1998, Voiceband Data Transmission Interface (Informative Reference)

[39] "Enhanced Services Deployment Guide" (1999). CableLabs®. Online at:
http://www.cablelabs.com/news/newsletter/SPECS/MayJune_SPECSTECH/tech.pgs/story2.html
[40] (Lipoff, 2002). "Security and Reliability of Hybrid-Fiber Co-ax Networks" Presentation to IGI Exec Telecom Briefing , October 30, 2002. eMail s.lipoff@ieee.org

[41] VoIP Availability and Reliability Model for the PacketCable Architecture Technical Report, Technical Report Number pkt-tr-voipar- v01-001128) at http://www.packetcable.com/ specifications/

[42] Wolf, Stephen (1990). "Features for Automated Quality Assessment of Digitally Transmitted Video", NTIA Report 90-264. Online at: http://www.its.bldrdoc.gov/pub/ntia-rpt/90-264/

[43] ITS Video Quality Research (2003). Web Site: http://www.its.bldrdoc.gov/n3/video/Default.htm

[44] tbd – audio format reference

[45] Steinmetz, Ralf (1996). Human Perception of Jitter and Media Synchronization. IEEE Journal on Selected Areas in Communications, Vol. 14, No. 1, January 1996

[46] Blakowski, Gerold, and Steinmetz, Ralf (1996). A Media Synchronization Survey: Reference Model, Specification and Case Studies.. IEEE Journal on Selected Areas in Communications, Vol. 14, No. 1, January 1996.

[47] ITU R BT.1359-1 (1998). Relative Timing of Sound and Vision for Broadcasting

[48] ATSC Implementation Subcommittee 191 "Relative Timing of Sound and Vision for Broadcast Operations", 26, June, 2003 (http://www.atsc.org/standards/is_191.pdf)

[49] Reference to McCann's law

[50] Reference to McCann's law

[51] Green et al. (2001). "Quality-of-service management for broadband residential video services", ELECTRONICS & COMMUNICATION ENGINEERING JOURNAL, 265-275,
DECEMBER 2001. Online at:
http://www.eee.bham.ac.uk/woolleysi/publications/ecej2001.pdf

[52] Bellcore, Wang et al. (1996). "Quality of Service (QoS) for Residential Broadband Video
Service", Global Telecommunications Conference, 1996. GLOBECOM '96. 'Communications: The Key to Global Prosperity ,Volume: 3 , 18-22 Nov. 1996 Pages:1949 - 1953 vol.3 .

[53] TR-110: DSLHome™ Reference Models for VoIP Configurations in the DSL Home, DSL Forum TR-110, September 2005.

[54] Schneiderman, B. (1998). Designing the User Interface: Strategies for Effective Human Computer Interaction. (3rd Ed). Reading, MA: Addison-Wesley.

[55] Gallaway, Glen R. (1981). Response Times to User Activities in Interactive Man/Machine Computer Systems. Proceedings of the Human Factors Society - 25th Annual Meeting, 754-758.

[56] Miller, Robert B. (1968), Response time in man-computer conversational transactions, Proceedings Spring Joint Conference 1968, 33, AFIPS Press, Montvale, NJ, 267-277.

[57] Youmans, D.M. (1983). The effects of system response time on users of interactive computer systems, IBM United Kingdom Laboratories, Hursley Park (January 1983).

[58] Willams, C.M. (1973). System response time: A study of users' tolerance, IBM Advanced Systems Development Division Technical Report, 17-272, Yorktown Heights NY (July 1973).

[59] Barber, Raymond E. and Lucus, Henry C. (1983). System Response Time, Operator Productivity, and Job Satisfaction. Communications of the ACM, Vol 26, No. 11, November 1983.

[60] Bickford, Peter (1998). Human Interface Online: Worth the Wait?, Netscape Communications, http://devedge.netscape.com/viewsource/bickford_wait.html

[61] Beerends, J.G & van der Gaast, S. (2004). Web Browsing Quality Modeling. ITU-T SG Contr 3.

[62] GameFlow - A Model for Evaluating Player Enjoyment in Games, Penelope Sweetser and Peta Wyeth, ACM Computers in Entertainment, Vol. 3, No. 3, July 2005. Article 3A.

[63] Järvinen, Aki, Satu Heliö and Frans Mäyrä (October 2002): "Communication and Community in Digital Entertainment Services". Prestudy Research Report. Hypermedia Laboratory Net Series 2.]

[64] ["Play and Playability as Key Concepts in New Media Studies", Julian Kücklich, Marie Curie Fellow, STeM Centre, Dublin City University]

[65] [Kline, Stephen, Nick Dyer-Witheford and Greig de Peuter (2003): Digital Play. The Interaction of Technology, Culture, and Marketing. Montreal and Kingston: McGill-Queen's University Press.]

[66] ["1500 Archers on a 28.8: Network Programming in Age of Empires and Beyond", Paul Bettner and Mark Terrano, Presented at GDC2001 - 03/22/2001 2:30p]

[67] ["OPScore, or Online Playability Score: A Metric for Playability of Online Games with Network Impairments", Ubicom Inc., March 2005],

[68] A.F. Wattimena, R.E. Kooij, J.M. van Vugt, O.K. Ahmed, "Predicting the perceived quality of a First Person Shooter: the Quake IV G-model", submitted for publication, NetGames 2006, 30-31 October 2006, Singapore

[69] [M. Dick, O. Wellnitz, L. Wolf (2005), "Analysis of factors affecting Players' Performance and Perception in Multiplayer Games", NetGames'05, Hawthorne (NY), U.S.A.]

[70] ["Aspects of Networking in Multiplayer Computer Games", Jouni Smed, Timo Kaukoranta, Harri Hakonen]

[71] Peter Quax, P., Monsieurs, P., Lamotte, W., De Vleeschauwer, D and Degrande, N. (2004). Objective and Subjective Evaluation of the Influence of Small Amounts of Delay and Jitter on a Recent First Person Shooter Game. SIGCOMM'04 Workshops, Aug.30+Sept.3, 2004, Portland, Oregon, USA.

[72] Beigbeder, T., Coughlan, R., Lusher, C., Plunkett, J., Agu, E. & Claypool,M. (2004). The Effects of Loss and Latency on User Performance in Unreal Tournament 2003. SIGCOMM'04 Workshops Aug. 30+Sept. 3, 2004, Portland, Oregon, USA.

[73] Pantel, L and Wolf, L.C. (2002). On the impact of delay on real-time multiplayer games. In Proc. of the 12th int. workshop on network and operating systems support for digital audio and video, pages 23–29. ACM Press, 2002.

[74] Armitage, G. (2001). Sensitivity of quake3 players to network latency. In *ACM SIGCOMM Internet Measurement Workshop 2001*, Berkeley, CA, USA, Nov. 2001.
[75] Armitage, G. (2003). An experimental estimation of latency sensitivity in multiplayer quake 3. In *11th IEEE Int.Conf. on Networks (ICON 2003),* Sydney, Australia, 2003.

[76] Schaefer, C., Enderes,T., Ritter,H & Zitterbart, M. (2002). Subjective Quality Assessment for Multiplayer Real-Time Games. NetGames2002, April 16-17, 2002, Braunschweig, Germany.

Voice Performance QoE and Implementation Guidelines– Informative References

Voice Service Quality Guidelines

TIA TSB-116 Voice Quality Recommendations for IP Telephony
TIA TSB-146 IP Telephony Support Emergency Calling Service
ITU G.114: Recommendations for delay on voice connections

- Quality contour used in the ITU E-Model to reflect impairment from delay on conversational (interactive) voice services.

TIA TSB-116.1, Voice Quality Recommendations for IP Telephony, Telecommunications Industry Association, 2001.

Essentials of Real-Time Networking – Provides details on voice QoE impacting parameters, and design targets in addition to tutorial information on QoS mechanisms. 2005, Nortel. Available online at:
http://www130.nortelnetworks.com/go/main.jsp?cscat=SUPPORT

ITU P.1010 Fundamental voice transmission objectives for VoIP terminals and gateways

- Provides performance requirements for narrow-band voice services for the whole range of packet-based gateways and terminals, including wireless and soft phones.

TIA-810-A Transmission Requirements for Narrowband Voice over IP and Voice over PCM Digital Wireline Telephones

- Performance requirements for digital terminals/end devices for IP and all digital networks. This standard sets out the performance expectations for all digital wireline terminals, regardless of the network type.

ITU G.1010 End-user Multimedia QoS Categories

- Provides guidance on the key factors that influence Quality of Service (QoS) from the perspective of the end-user.

ITU Y.1540 Internet protocol data communication service - IP packet transfer and availability performance parameters

ITU Y.1541 Network performance objectives for IP-based services

Standards Governing Prediction and Measurement of Voice Performance

ITU G.107 The E-Model, a computational model for use in transmission planning

ITU G.113 Transmission Impairments due to Speech Processing

- Provides guidance regarding transmission impairments introduced by digital speech processing systems for use in conjunction with the transmission planning approach described in G.107, G.108 and G.109. The Impairment Factor method, used by the E-model of G.107, replaces the Quantization Distortion (QDU) system for non-G.711 systems. Appendix I contains updated Impairment Factor values for various digital processing systems; Appendix II contains guidance on how an Advantage Factor can be used to reflect the variation in user expectation of quality for different communications systems (e.g. mobile).

ITU P.862 Perceptual evaluation of speech quality (PESQ), an objective method for end-to-end speech quality assessment of narrowband telephone networks and speech codecs
- Defines an algorithm intended to estimate the subjective quality of narrowband voice signals. More on P.862 and PESQ.

ITU P.862.1 Mapping function for transforming P.862 raw result scores to MOS-LQO
- Provides a standard method to transform raw PESQ scores into a MOS estimate. Replaces PESQ-LQ transformation defined by Psytechnics.

ITU P.862.3 Application Guidelines for the use of PESQ

ITU-T Recommendation P.800, Methods for subjective determination of transmission quality, Geneva: ITU-T, 1996.

ITU-T Recommendation P.800.1, Mean Opinion Score (MOS) terminology, Geneva: ITU-T, 2003.

ITU-T Recommendation P.830, Subjective performance assessment of telephone-band and wideband digital codecs, Geneva: ITU-T, 1996.

ITU-T Recommendation P.861, Objective quality measurement of telephone-band (300-3400 Hz) speech codecs, Geneva: ITU-T, 1998 (withdrawn).

ITU-T Recommendation P.862, Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrowband telephone networks and speech codecs, Geneva: ITU-T, 2001.

ITU-T Recommendation P.862.1, Mapping function for transforming P.862 raw result scores to MOS-LQO, Geneva: ITU-T, 2003.

ITU-T Recommendation PESQ-LQ, A.W. Rix, Comparison between subjective listening quality and P.862 PESQ score, Psytechnics, September 2003.

Advisory & Implementation Guidance Documents for IP Telephony

TIA-912-A Voice Gateway Transmission Requirements (including applicable loss plan)
TIA-464-C Requirements for PBX Switching Equipment (including applicable loss plan)

T1.508 Loss Plan for Digital Networks

- The loss plan is important for control of voice signal level and control of echo within the network, as well as across network boundaries. Transmission recommendations for connections between network interfaces (NIs) that are all-digital (NI-to-NI), or that are all-digital from digital end office-to-digital end office) with digital or analog access lines from appropriate end-user terminals or interconnecting networks. Includes new recommendation for loss applied to local analog access lines in TDM networks.

ITU G.131 Talker Echo and its Control

- Attenuation requirements for controlling echo

ITU G.168 Digital network echo cancellers

- Performance requirements for digital echo cancellers.

ITU G.165 Echo cancellers

ITU-T Recommendation G.711, Pulse code modulation (PCM) of voice frequencies, Geneva: ITU-T, 1988.

ITU-T Recommendation G.726, 40, 32, 24, 16 kbit/s Adaptive Differential Pulse Code Modulation (ADPCM), (includes Annex A: Extensions of Recommendation G.726 for Use with Uniform-Quantized Input and Output-General Aspects of Digital Transmission Systems), Geneva: ITU-T, 1990.

ITU-T Recommendation G.729, Coding of Speech at 8 kbit/s Using Conjugate-Structure Algebraic-Code-Excited Linear-Prediction (CSACELP),

Rec. G.729, (includes Annex A, Reduced Complexity 8 kbit/s CS-ACELP Speech Codec, and Annex B, Silence Compression Scheme for G.729 Optimized for Terminals Conforming to Recommendation V.70), Geneva: ITU-T, 1996.

T1.508-2003, Loss plan for digital networks, Washington, DC: American National Standards Institute, Committee T1, 2003.

TIA-810-A9, Transmission requirements for narrowband Voice over IP and Voice over PCM digital wireline telephones.

G.113, Transmission Impairments due to Speech Processing, (includes Appendix I, Provisional planning values for the equipment impairment factor Ie and packet-loss robustness factor Bpl.) Geneva: ITU-T, 2001.