ETSI TR 102 805-3 V1.1.1 (2010-04)

Technical Report

User Group;
End-to-end QoS management at the Network Interfaces;
Part 3: QoS informational structure

Reference

DTR/USER-00029-3

Keywords

QoS, interface

ETSI

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

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

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

Important notice

Individual copies of the present document can be downloaded from:
http://www.etsi.org

The present document may be made available in more than one electronic version or in print. In any case of existing or perceived difference in contents between such versions, the reference version is the Portable Document Format (PDF). In case of dispute, the reference shall be the printing on ETSI printers of the PDF version kept on a specific network drive within ETSI Secretariat.

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

If you find errors in the present document, please send your comment to one of the following services:
http://portal.etsi.org/chaircor/ETSI_support.asp

Copyright Notification

Contents

Intellectual Property Rights

IPRs essential or potentially essential to the present document may have been declared to ETSI. The information pertaining to these essential IPRs, if any, is publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (http://webapp.etsi.org/IPR/home.asp).

Pursuant to the ETSI IPR Policy, no investigation, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Foreword

This Technical Report (TR) has been produced by ETSI User Group (USER).

The present document is part 3 of a multi-part deliverable covering the End-to-end QoS management at the Network Interfaces, as identified below:

Part 1: "User's E2E QoS - Analysis of the NGN interfaces (user case)";

Part 2: "Control and management planes solution: QoS continuity";

Part 3: "QoS informational structure".

Introduction

The network and service QoS management is a multidimensional problem. In the methodology described in parts 1 [i.1] and 2 [i.2], five dimensions were identified through the analysis carried out: the **information dimension** which represents the entire system by the structured data; the **organizational dimension** which defines the management/control relationship between different entities; the **functional dimension** which defines a set of functions related to management/control activities, the **architectural dimension** which describes the structure of management entities and their related interfaces for information exchange, and **the protocol dimension** which defines the means for transmitting the management information. Among these five dimensions, the information dimension is the basic building block for the other dimensions of the End-to-End (E2E) QoS management and control. This structure of the information dimension provides a generic information image of the components in the user's system and a description of any ambient resource applied to any architecture and system technology for the QoS management. The present document focuses on the QoS informational structure.

1 Scope

The present document provides a description of QoS criteria related to the different visibility levels (Equipment, Network, Service) and related profiles involved in each step of the service lifecycle. A set of user preferences and an information structure are defined in the user-centric profile for personalisation purposes.

2 References

References are either specific (identified by date of publication and/or edition number or version number) or non-specific.

- For a specific reference, subsequent revisions do not apply.
- Non-specific reference may be made only to a complete document or a part thereof and only in the following cases:
  - if it is accepted that it will be possible to use all future changes of the referenced document for the purposes of the referring document;
  - for informative references.

Referenced documents which are not found to be publicly available in the expected location might be found at http://docbox.etsi.org/Reference.

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

2.1 Normative references

The following referenced documents are indispensable for the application of the present document. For dated references, only the edition cited applies. For non-specific references, the latest edition of the referenced document (including any amendments) applies.

Not applicable.

2.2 Informative references

The following referenced documents are not essential to the use of the present document but they assist the user with regard to a particular subject area. For non-specific references, the latest version of the referenced document (including any amendments) applies.

[i.1] ETSI TR 102 805-1: "User Group; End-to-end QoS management at the Network Interfaces; Part 1: User's E2E QoS - Analysis of the NGN interfaces (user case)".

[i.2] Noëmie Simoni, Simon Znaty (1997): "Gestion de réseau et de service: similitude des concepts, spécificité des solutions".

[i.3] ETSI ES 202 746: "Human Factors (HF); Personalization and User Profile Management; User Profile Preferences and Information".

[i.4] IETF RFC 1633: "Integrated Services in the Internet Architecture: an Overview".

[i.5] IETF RFC 2474: "Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers".

[i.6] IETF RFC 2475: "Architecture for Differentiated Services".

[i.7] ETSI TR 102 805-2: "User Group; End-to-end QoS management at the Network Interfaces; Part 2: Control and management planes solution - QoS continuity".

3 Definitions and abbreviations

3.1 Definitions

For the purposes of the present document, the following terms and definitions apply:

AmbientGrid: information inference (AmbientGrid) based on the profiles' matching, to structure with grid covering the needed user centric environment

class of service: way of traffic management in the network by grouping similar types of traffic and treating them as its own level of service priority

Diffserv: DiffServ networks classify packets into one of a small number of aggregated flows or 'classes', based on the DiffServ codepoint (DSCP) in the packet's IP header

NOTE: This is known as behaviour aggregate (BA) classification (RFC 2475 [i.6]). At each DiffServ router, packets are subjected to a 'per-hop behaviour' (PHB), which is invoked by the DSCP (RFC 2474 [i.5]).

infosphere: decisional knowledge base managing, in the real time, all the personalization and ambient environment information

infoware: knowledge base that covers the different visibility levels and acts by inference

IntServ: the integrated services architecture RFC 1633 [i.4] defined a set of extensions to the traditional best effort model of the Internet with the goal of allowing end-to-end QoS to be provided to applications

NOTE: One of the key components of the architecture is a set of service; the current set of services consists of the controlled load and guaranteed services. The architecture assumes that some explicit setup mechanism is used to convey information to routers so that they can provide requested services to flows that require them. While RSVP is the most widely known example of such a setup mechanism, the IntServ architecture is designed to accommodate other mechanisms.

QoS classification: definition of class priority for QoS by describing traffic condition or performance parameters

service mobility: ability to consistently provide services to the end-user, to maintain the expected QoS, at the system's initiative, regardless of the end-user's location, terminals, or networks

NOTE: To maintain the service continuity, the session mobility is used.

terminal mobility: use of a mobile device while moving across the same or different networks and having access to the same set of subscribed services

user mobility: ability for a subscriber to move to different physical locations and be able to use one or more devices connected to one or more access networks to gain access to their services without interruption

user-centric session: period of communication between one end-user and another or other end-users or servers characterized by a starting time and a termination time, including setting up the relation of the end-user equipment, access network, core network and services invoked during this period

userware: innovative user centric middleware (Userware) enhancing the seamless feasibility along with the location and activity, personalization and user's ambient contexts

3.2 Abbreviations

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

BA	Behaviour Aggregate
CPU	Central Processing Unit
DiffServ	Differentiated services (IETF)
DSCP	Differentiated Services CodePoint
E2E	End-to-End
ETSI	European Telecommunications Standards Institute

GGSN	Gateway GPRS Service Node
HSS	Home Subscriber Server
IETF	Internet Engineering Task Force
IntServ	Integrated Services (IETF)
LM	Local Machine
MIPS	Millions of Instructions Per Second
MMS	Multimedia Messaging Service
NGN	Next Generation Network
NLN	Node-Link-Network
QoS	Quality of Service
RS	Remote Server
SGSN	Serving GPRS Support Node
SLA	Service Level Agreement
SMS	Short Message Service
UMTS	Universal Mobile Telecommunications Systems
VPCN	Virtual Private Connectivity Network
VPEN	Virtual Private Equipment Network
VPSN	Virtual Private Service Network

4 QoS and profiles

In this clause, the instantiations of each visibility level (Equipment, Network and Service) are described according to the QoS criteria defined in [i.1] (clause 4.1). The different QoS information involved in each step of the service's lifecycle (Figure 1) are then presented: resource profile with QoS capability which is involved in the E2E conception phase (clause 4.2), resource usage profile which is involved in the deployment phase (clause 4.3) and finally the QoS values which are involved in the exploitation phase (clause 4.4).

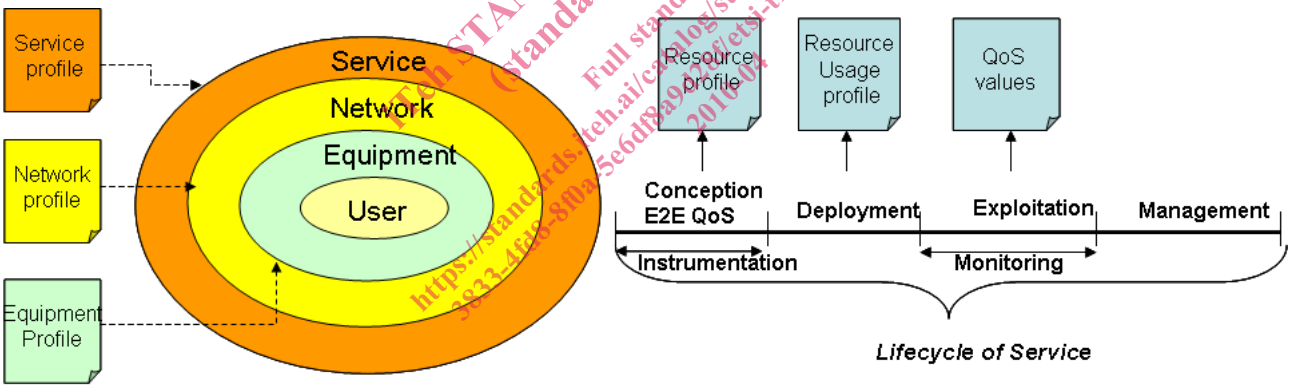

Figure 1: QoS and profiles

4.1 QoS criteria related to visibility levels

The QoS of the four actors (Equipment, Network, Service and User) participating in the E2E session determines the overall E2E session's QoS (Figure 2). Each actor has responsibilities in the achieved QoS [i.2].

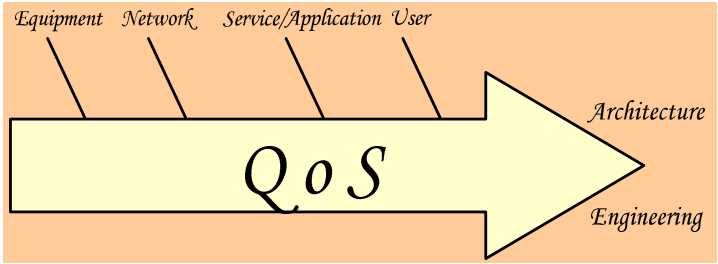

Figure 2: End-to-end QoS

In order to monitor the component behaviour in a heterogeneous environment, it is necessary that a homogeneous expression of its QoS be available to evaluate the E2E behaviour during the whole session.

Therefore, a unified QoS model (four criteria: ***availability, delay, fidelity*** and ***capability***) is applied to all components. These criteria can be applied to any QoS classification (Diffserv, Interserv, etc) and can also be measured easily according to specific parameters:

- Availability represents the aptitude of a service, network or equipment element to be accessed at a certain moment. The availability depends on the demands and the contractual conditions of the environment and time. It indicates the accessibility rate of the nodes and the logical links which have also been defined in the Meta model.
- Fidelity represents the aptitude of a service, network or equipment element to work without deteriorating the treated information with respect to the demands and the contractual conditions. It indicates the involuntary modification rate of the information concerning the nodes and the links during their treatment.
- Delay represents the aptitude of a service, network or equipment element to comply with the specific times requested in the demands and the contractual conditions. It indicates the treatment's duration of the nodes and the mean delay of the link transit.
- Capability represents the aptitude of a service, network or equipment element, which should have the necessary means to realize its task according to the demands and the contractual conditions. It indicates the maximum charge of the nodes and the usable bandwidth of the links.

The instantiations of the QoS model at each level are shown in the tables below. They demonstrate the feasibility of the measurement to evaluate the QoS in the performance assessment.

For the equipment visibility level, QoS parameters depend on the characteristics of hardware hosting the different software. Table 1 shows the instantiation of the QoS criteria at equipment visibility level.

Table 1: Instantiation of the QoS criteria at equipment level

QoS Criteria	Equipment		
	Node	Link	Network
Availability	Memory size	Linking (i.e. cable, fibre) availability	Availability resulting of the VPEN
Fidelity	Error rate in server/router/device	Error rate in link	Error rate of the VPEN
Delay	Handling time in server/router/device. Service time provided by the equipment depending on battery life	Propagation delay time	E2E delay time
Capability	MIPS of CPU	Standard bandwidth of link depending on technology	Standard bits per second
MIPS: Millions of instructions per second.			

For the network visibility level, the QoS parameters depend on the functional and procedural means for the data units' transit in the network link. Table 2 shows the instantiation of the QoS criteria at the network visibility level.

Table 2: Instantiation of the QoS criteria at network level

QoS Criteria	Network		
	Node	Link	Network
Availability	Accessibility rate of protocol process	Rejection rate	Accessibility rate of the VPCN
Fidelity	Packet error rate Duplication of packet	Packet error rate during routing	Error rate of the VPCN
Delay	Handling and information transit time of protocol machine	Delay for establishing the network connection + transit time	E2E transit time in network
Capability	Number of packet handled per second	Routed packet number handled per second	Delivered packet number handled per second

For the service visibility level, the QoS parameters depend on the functional and procedural means for establishing the service process through the associated link. Table 3 shows the instantiation of the QoS criteria at the service visibility level.

Table 3: Instantiation of the QoS criteria at service level

QoS Criteria	Service		
	Node	Link	Network
Availability	Accessibility rate of the Service Element (SE)	Association rupture rate	Accessibility rate of the VPSN
Fidelity	Error rate of the Service Element.	Error rate of message	Error rate of the VPSN
Delay	Handling and information transit time in the Service Element	Delay for establishing the association between the Service Elements	E2E Response time
Capability	Number of messages treated per second	Number of messages exchanged per second	Number of messages delivered per second

4.2 Resources profiles

Resource profiles are involved in the phase of the E2E conception. They are the basic profiles representing all the resources in a structured and uniform format. They contain the characteristics and definition of the resources at the visibility level of the service, network and terminal with the QoS capability. Resource profile is independent of any execution environment (anyway, in an optimization context) and is published by the resource provider.

The Equipment profile, Network profile and Service profile are instantiated from the resource profile. In the phase of E2E conception, in the service's lifecycle, each profile indicates the maximum QoS capabilities at its own level that a resource can provide with respect to the four criteria. For example, in Figure 3, a service profile is instantiated from the resource profile where the QoS capabilities are detailed according to the four QoS criteria.


```
<?xml version="1.0" encoding="UTF-8" ?>
- <ServiceProfile>
    <version>1.0.0.0</version>
  - <Identity>
      <ServiceName>VoD-SFR</ServiceName>
      <ServiceType>streaming</ServiceType>
    </Identity>
  - <QoSCapabilities>
      <Availability>0.998888</Availability>
      <Fidelity>0.998888</Fidelity>
      <Delay>3000ms</Delay>
      <Capability> 0.5M </Capability>
    </QoSCapabilities>
  </ServiceProfile>
```


Figure 3: Service resource profile

4.3 Resource usage profiles

Resource usage profiles are involved in the service deployment phase. Resource usage profiles (of equipment, network and service) provide a uniform and well structured format for each component's possibilities and constraints. Moreover, the QoS contract (Demanded QoS and Offered QoS with specific values) is integrated in each component (Figure 4), that allows selecting the components in a user centric context; i.e. the selected components will respond to the request and adapt the selection according to the user's context.

Meanwhile, the Resource Usage Profile is related to the execution environment of a resource. A component should be supported by the appropriate Demanded QoS from the lower level in order to offer an expected QoS to the upper layer (Figure 4). Demanded QoS will be compared with the Offered QoS in order to have a service provided with the right QoS. Therefore, the QoS cross-layer chaining will be transparent to the user.

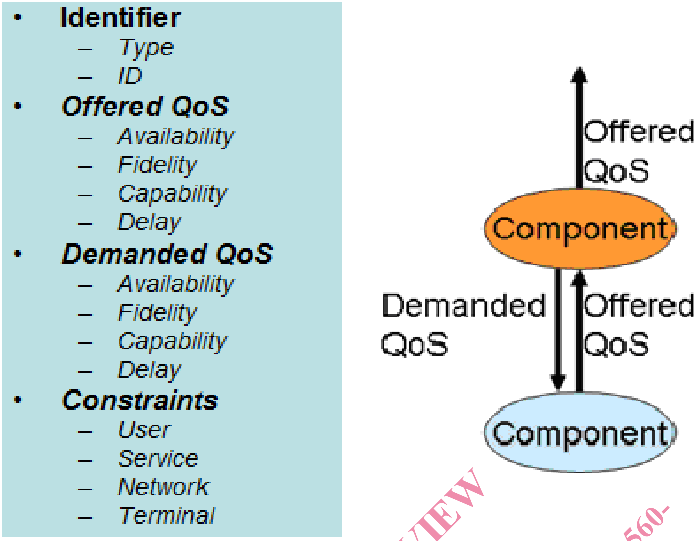

Figure 4: Resource usage profile structure

A Usage profile contains the following information:

Identifier: this field defines the resource type and allows identifying each resource without any addressing restriction by using a unique identifier.

Offered QoS: this field indicates the expected QoS of the described component resource.

Demanded QoS: this field represents the QoS needed in order to allow the described resource to work correctly.

Constraints: this field determines the restrictions of each level.

An example of service usage profile (end-to-end QoS vision) is shown in Figure 5.


```
<?xml version="1.0" encoding="UTF-8" ?>
- <ServiceUsageProfile>
    <version>1.0.0.0</version>
  - <Identity>
      <ServiceName>VoD-SFR</ServiceName>
      <ServiceType>streaming</ServiceType>
    </Identity>
  - <OfferedQoS>
      <Availability>0.998888</Availability>
      <Fidelity>0.998888</Fidelity>
      <Delay>3600ms</Delay>
    - <Capability>
        <UserMax>300</UserMax>
      </Capability>
    </OfferedQoS>
  - <DemandedQoS>
      <Availability>0.998888</Availability>
      <Fidelity>0.998888</Fidelity>
      <Delay>3000ms</Delay>
      <Capability>0.5M</Capability>
    </DemandedQoS>
  - <Constraints>
    - <User>
        <Group>AnyBody</Group>
        <Langage>English</Langage>
        <AccessCost>"0.20/m"</AccessCost>
      </User>
    - <Service>
        <State />
        <Preference />
      </Service>
    - <Network>
        <ProtocolRelated>IPV6</ProtocolRelated>
        <ConnectionExtern>TRUE</ConnectionExtern>
      </Network>
    - <Equipement>
        <Processor>2.5GHz</Processor>
        <Memory>2Go</Memory>
        <Application>MediaPlayer</Application>
      </Equipement>
    </Constraints>
  </ServiceUsageProfile>
```


Figure 5: Service usage profile