ETSI TS 102 250-8-1 V1.1.1 (2018-12)

TECHNICAL SPECIFICATION

Speech and multimedia Transmission Quality (STQ);
QoS aspects for popular services in mobile networks;
Part 8: Formalized definition of Quality of Service parameters and their computation;
Sub-part 1: General aspects and terminology

Reference
DTS/STQ-00218m

Keywords
3G, GSM, network, QoS, service, speech

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

Contents

Intellectual Property Rights

Essential patents

IPRs essential or potentially essential to normative deliverables 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 (https://ipr.etsi.org/).

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.

Trademarks

The present document may include trademarks and/or tradenames which are asserted and/or registered by their owners. ETSI claims no ownership of these except for any which are indicated as being the property of ETSI, and conveys no right to use or reproduce any trademark and/or tradename. Mention of those trademarks in the present document does not constitute an endorsement by ETSI of products, services or organizations associated with those trademarks.

Foreword

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

The present document is part 8, sub-part 1 of a multi-part deliverable. Full details of the entire series can be found in part 1 [36].

Modal verbs terminology

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

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

Introduction

The definition of quality of service (QoS) parameters and their computation, as in ETSI TS 102 250-2 [i.1], involves a couple of formal aspects which require special attention but have not been addressed systematically so far by the ETSI TS 102 250 document series. These aspects include:

- Introduction and consistent usage of appropriate symbols in mathematical equations according to the ETSI Drafting Rules.
- Clear distinction between measurable quantities and associated statistical aggregations.
- Unambiguous naming of trigger events which occur in the definitions of several different QoS parameters.
- Systematic introduction and definition of appropriate points of observation which form the link between harmonized, abstract QoS parameter definitions and their application to specific contexts, e.g. particular types of access networks.

Over the years, application and maintenance of ETSI TS 102 250-2 [i.1] has revealed some deficiencies in this respect, which are not easily solvable via ordinary change requests, but would require structural changes of ETSI TS 102 250-2 [i.1]. Therefore, a new part 8 of the ETSI TS 102 250 series is introduced, of which the present document represents the first sub-part.

The present document treats the aspects listed above in a systematic way and specifies general requirements for a formalized definition of QoS parameters. Subsequent sub-parts of ETSI TS 102 250-8 are envisaged to contain QoS parameter definitions in accordance with the present document. These definitions will then replace, and thus obsolete, corresponding parts of ETSI TS 102 250-2 [i.1].

Clause 4 introduces the basic terminology used in a formalized treatment of QoS parameters: After clarifying the difference between measurable quantities and statistical averages, the concept of points of observation is explained in detail. Afterwards, the notion of abstract and technical events is introduced and used to specify a methodology for deriving harmonized QoS parameters in two stages. Finally, a formalized, event-based characterization of the term "transaction" is given, which proves to be useful when modelling complex scenarios of service usage.

Clause 5 specifies generic symbols and formulas for types of quantities frequently occurring in QoS parameter definitions, like event indicators, time durations or data rates. In addition, formally correct definitions for commonly used sample averages are given.

A systematic approach for deriving QoS parameters for non-trivial services, which may be subdivided into different "phases", is outlined in clause 6.

Finally, clause 7 defines user-oriented points of observation for different types of access networks. The exact locations of these points are illustrated in protocol stack overview figures, and reference is made to the corresponding protocol specifications.

Two informative annexes complement the main text of the present document:

- Annex A gives some guidance how to apply the rather formal definitions of clause 5 in practice by outlining a contrived, but complete, example.
- Annex B contains guidelines for writing specifications according to the present document. These guidelines include suggestions for structuring the respective document (for example, one of the intended further sub-parts of ETSI TS 102 250-8 [i.3]) as well as templates for particular parts of such specifications.

1 Scope

The present document defines a consistent terminology for the formalized definition of quality of service (QoS) parameters and their computation. In doing so, it carefully distinguishes between measurable quantities and associated statistical aggregation formulas. For frequently used types of quantities and aggregations, generic symbols and calculation formulas are specified.

For a large class of QoS parameters, which are basically determined by the observation of events, the present document specifies a methodology for constructing definitions in two stages, using the concept of abstract events and their representation by technical events at well-defined points of observation.

In addition, a framework is presented which indicates how QoS parameters may be derived in a generic way via sub-dividing the usage of a service into meaningful sub-transactions, called "phases".

2 References

2.1 Normative references

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

Referenced documents which are not found to be publicly available in the expected location might be found at https://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.

The following referenced documents are necessary for the application of the present document.

[1] ETSI TS 102 250-6: "Speech Processing, Transmission and Quality Aspects (STQ); QoS aspects for popular services in GSM and 3G networks; Part 6: Post processing and statistical methods".

[2] ETSI TS 102 250-7: "Speech and multimedia Transmission Quality (STQ); QoS aspects for popular services in GSM and 3G networks; Part 7: Network based Quality of Service measurements".

[3] ETSI TS 124 002: "Digital cellular telecommunications system (Phase 2+) (GSM); Universal Mobile Telecommunications System (UMTS); LTE; GSM - UMTS Public Land Mobile Network (PLMN) Access Reference Configuration (3GPP TS 24.002)".

[4] ETSI TS 143 051: "Digital cellular telecommunications system (Phase 2+); Overall description; Stage 2 (3GPP TS 43.051)".

[5] ETSI TS 123 060: "Digital cellular telecommunications system (Phase 2+) (GSM); Universal Mobile Telecommunications System (UMTS); General Packet Radio Service (GPRS); Service description; Stage 2 (3GPP TS 23.060)".

[6] ETSI TS 123 401: "LTE; General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access (3GPP TS 23.401)".

[7] ETSI TS 127 007: "Digital cellular telecommunications system (Phase 2+) (GSM); Universal Mobile Telecommunications System (UMTS); LTE; AT command set for User Equipment (UE) (3GPP TS 27.007)".

[8] ETSI TS 127 060: "Digital cellular telecommunications system (Phase 2+) (GSM); Universal Mobile Telecommunications System (UMTS); LTE; Packet domain; Mobile Station (MS) supporting Packet Switched services (3GPP TS 27.060)".

[9] IETF RFC 3501: "Internet message access protocol - version 4rev1".

[10] IETF RFC 1939: "Post Office Protocol - Version 3".

[11] IETF RFC 5321: "Simple Mail Transfer Protocol".

[12] IEEE Std 802.11™: " IEEE Standard for Information technology - Telecommunications and information exchange between systems Local and metropolitan area networks - Specific requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications".

[13] IETF RFC 1034: "Domain names - concepts and facilities".

[14] IETF RFC 1035: "Domain names - implementation and specification".

[15] IETF RFC 959: "File Transfer Protocol".

[16] IETF RFC 3261: "SIP: Session Initiation Protocol".

[17] IETF RFC 3550: "RTP: A Transport Protocol for Real-Time Applications".

[18] IETF RFC 3711: "The Secure Real-time Transport Protocol (SRTP)".

[19] IETF RFC 7826: "Real-Time Streaming Protocol Version 2.0".

[20] IETF RFC 7230: "Hypertext Transfer Protocol (HTTP/1.1): Message Syntax and Routing".

[21] IETF RFC 2131: "Dynamic Host Configuration Protocol".

[22] ETSI EN 300 392-5: "Terrestrial Trunked Radio (TETRA); Voice plus Data (V+D) and Direct Mode Operation (DMO); Part 5: Peripheral Equipment Interface (PEI)".

[23] ETSI TS 144 001: "Digital cellular telecommunications system (Phase 2+); Mobile Station - Base Station System (MS - BSS) interface; General aspects and principles (3GPP TS 44.001)".

[24] ETSI TS 144 004: "Digital cellular telecommunications system (Phase 2+) (GSM); GSM/EDGE Layer 1; General Requirements (3GPP TS 44.004)".

[25] ETSI TS 144 005: "Digital cellular telecommunications system (Phase 2+) (GSM); GSM/EDGE Data Link (DL) Layer; General aspects (3GPP TS 44.005)".

[26] ETSI TS 145 008: "Digital cellular telecommunications system (Phase 2+) (GSM); GSM/EDGE Radio subsystem link control (3GPP TS 45.008)".

[27] ETSI TS 124 007: "Digital cellular telecommunications system (Phase 2+) (GSM); Universal Mobile Telecommunications System (UMTS); LTE; Mobile radio interface signalling layer 3; General Aspects (3GPP TS 24.007)".

[28] ETSI EN 300 392-1: "Terrestrial Trunked Radio (TETRA); Voice plus Data (V+D); Part 1: General network design".

[29] ETSI EN 300 392-2: "Terrestrial Trunked Radio (TETRA); Voice plus Data (V+D); Part 2: Air Interface (AI)".

[30] ETSI TS 125 301: "Universal Mobile Telecommunications System (UMTS); Radio interface protocol architecture (3GPP TS 25.301)".

[31] ETSI TS 124 301: "Universal Mobile Telecommunications System (UMTS); LTE; Non-Access-Stratum (NAS) protocol for Evolved Packet System (EPS); Stage 3 (3GPP TS 24.301)".

[32] ETSI TS 136 331: "LTE; Evolved Universal Terrestrial Radio Access (E-UTRA); Radio Resource Control (RRC); Protocol specification (3GPP TS 36.331)".

[33] Recommendation ITU-T I.411: "ISDN user-network interfaces - Reference configurations".

[34] Recommendation ITU-T I.412: "ISDN user-network interfaces - Interface structures and access capabilities".

[35] Recommendation ITU-T Q.931: "ISDN user-network interface layer 3 specification for basic call control".

[36] ETSI TS 102 250-1: "Speech and multimedia Transmission Quality (STQ); QoS aspects for popular services in mobile networks; Part 1: Assessment of Quality of Service".

2.2 Informative references

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

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

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

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

[i.2] Recommendation ITU-T X.290: "OSI conformance testing methodology and framework for protocol Recommendations for ITU-T applications - General concepts".

[i.3] ETSI TS 102 250-8 (all sub-parts): "Speech and multimedia Transmission Quality (STQ); QoS aspects for popular services in mobile networks; Part 8: Formalized definition of Quality of Service parameters and their computation".

3 Definitions of terms, symbols and abbreviations

3.1 Terms

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

abstract event: abstractly formulated proposition, which can be said to be either true or false

point of observation: combination of a particular reference point with a particular protocol layer or service access point at that reference point

technical event: observation that a well-defined piece of data has been either sent or received

3.2 Symbols

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

$\coloneqq$	"is defined by" operator stating that the quantity on its left-hand side is defined by the expression on its right-hand side
d_T	data rate of a data transfer transaction T
$\widetilde{d_T}$	mean data rate of a data transfer transaction T
Δt_T	duration of a transaction T
$\widetilde{\Delta t_T}$	mean duration of a transaction T
ΔD_T	amount of data transferred in a transaction T
$E_{\mathrm{end},T}$	event defining the end of a transaction T
$E_{\mathrm{start},T}$	event defining the start of a transaction T
$E_{\mathrm{success},T}$	success condition of a transaction T
I_{A}	indicator of an event A
$n(A)$	number of measurement cycles where an event A has occurred

Q	measurable quantity
$\bar{Q}$	sample average of a measurable quantity
$\overline{Q\|B}$	sample average of a measurable quantity Q under the condition that an event B has occurred
$r(A)$	sample ratio of an event A
$r(A\|B)$	sample ratio of an event A under the condition that an event B has occurred
$r_{\mathrm{fail},T}$	failure ratio of a transaction T
$r_{\mathrm{success},T}$	success ratio of a transaction T
T	transaction
t_{A}	time of occurrence of an event A

3.3 Abbreviations

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

API	Application Programming Interface
AT	ATtention command prefix
CLI	Calling Line Identification
CMCE	Circuit Mode Control Entity
CR	Change Request
CS	Circuit Switched
DHCP	Dynamic Host Configuration Protocol
EDGE	Enhanced Data rates for GSM Evolution
FTP	File Transfer Protocol
GERAN	GSM/EDGE Radio Access Network
GSM	Global System for Mobile communications
HTTP	Hyper Text Transport Protocol
ISDN	Integrated Services Digital Network
LLC	Logical Link Control
MAC	Medium Access Control
MLME	MAC subLayer Management Entity
MMI	Man Machine Interface
MMS	Multimedia Messaging Service
MOS	Mean Opinion Score
MS	Mobile Station
MT	Mobile Termination
n.a.	not available
PCO	Point of Control and Observation
PO	Point of Observation
PPP	Point-to-Point Protocol
PS	Packed Switched
QoS	Quality of Service
RT	Reference point between TE and MT in TETRA
RTP	Realtime Transport Protocol
RTCP	Realtime Transport Control Protocol
RTSP	Real Time Streaming Protocol
SIM	Subscriber Identity Module
SIP	Session Initiation Protocol
SME	Station Management Entity
SNDCP	SubNetwork Dependant Convergence Protocol
TA	Terminal Adapter
TE	Terminal Equipment
TETRA	TErrestrial Trunked RAdio
UE	User Equipment
UMTS	Universal Mobile Telecommunications Service
UNI	User Network Interface
UTRAN	Universal Terrestrial Radio Access Network
XML	eXtensible Markup Language

4 QoS parameter basics

4.1 Measurable quantities and sample averages

The document series ETSI TS 102 250-8 [i.3] defines QoS parameters as quantitative measures for particular quality aspects of popular services in mobile networks. The definition of such a quantitative measure shall consist of two parts:

1) A **measurable quantity** Q associated with the respective QoS parameter shall be defined. This means, an explicit equation or algorithm shall be given which represents a unique instruction how to determine the measured value of the respective QoS parameter from more elementary observation data (like events, speech samples, etc.) in a single measurement.

2) An **aggregation formula or algorithm** shall be specified which maps a set $\{q_1, \dots, q_N\}$ of results of repeated measurements of the same quantity Q to a single value. This aggregated value then provides the desired **measure** associated with the respective QoS parameter.

The rationale behind this approach is the following: Mobile networks are complex dynamical systems, whose state undergoes non-systematic fluctuations on small time scales (in fact, the same is true not only for mobile networks, but for any telecommunication network). Possible origins of these fluctuations are varying radio conditions, varying utilization by different users and many other factors.

Due to this fluctuating behaviour, a single measurement of any quantity representing a QoS aspect (e.g. the time needed to set up a particular telephony call) usually does not yet yield a meaningful measure for this QoS aspect: Repeated measurements of the same quantity will in general yield different results which may exhibit considerable variations, even while there is no systematic change in the network conditions.

NOTE: These variations have their origin in the incomplete control over the state of the network (it is practically impossible to know all the parameters influencing this state) and should not be confused with uncertainties introduced by the measurement equipment - the variations would be present even if measured with an ideal, error-free measurement device. Of course, any real measurement system will add yet another contribution (to be kept as small as possible) to the variability of the results.

In order to deal with the fluctuating nature of a mobile network, a statistical approach is taken: The same quantity is measured repeatedly, and statistical aggregation of the single results will then provide the desired measure. Loosely speaking, statistical aggregation will "cancel out" non-systematic fluctuations of the single measurement results in favour of statistically relevant information, like mean or median values, variances, percentiles, etc.

A frequently used aggregation of a set of N measurement results $\{q_1, \dots, q_N\}$ for the same quantity Q is the arithmetic mean of the result set, also called the "sample average":

$$\bar{Q} := \frac{1}{N}\sum_{i=1}^{N} q_i \tag{4.1}$$

For specific purposes, other statistical aggregations, like median, quantiles, etc. may be used as well (for details please refer to ETSI TS 102 250-6 [1]).

The reader should distinguish between the **measurable quantity** Q associated with a QoS parameter on the one hand, and the **sample average** $\bar{Q}$ of this quantity on the other hand: The former comprises the elementary definition of the QoS parameter by a measurement instruction, while the latter represents the proper **measure** (involving statistical aggregation) associated with this QoS parameter.

Nevertheless, for the sake of simplicity, the term "QoS parameter" will be used in the present document with either meaning whenever it is clear from the context that no ambiguity can arise.

Depending on the specific quality aspects to be reflected by a particular QoS parameter, the formula or algorithm defining the corresponding measurable quantity Q may already involve aggregation itself (from yet more elementary measurement data). However, this kind of aggregation still needs to be distinguished from the one employed to get sample averages. In other words, the definition of the measurable quantity determines what is regarded to be a "sample" in a particular context. Several alternatives may be possible, the choice being a matter of convenience.

Consider for example speech quality measures for telephony calls: In addition to a "speech quality on sample basis", where the measurable quantity basically consists of a mean opinion score (MOS) for a single speech sample, another QoS parameter "speech quality on call basis" may be introduced, where the definition of the measurable quantity involves an aggregation over all speech samples in a particular call. In the former case, the sample space for statistical calculations consists of as many values as there have been speech samples in all calls considered, whereas in the latter case the number of statistical samples equals the number of calls.

4.2 Points of observation

To avoid ambiguities when defining QoS parameters in terms of measurable quantities, the definition of **what** is to be measured shall be complemented with a precise specification of **where** (in the functional architecture of the telecommunication network) the respective quantity shall be observed. Conversely, the specific representation of an (abstractly characterized) quantity - in other words its **final** definition - for a particular context usually depends on the point where it is to be measured in that context.

EXAMPLE: Consider the time needed to set up a phone call. This (abstract) quantity makes sense in different network technologies (e.g. ISDN, 2G/3G mobile networks, VoIP, VoLTE), and even within one technology one may wish to measure it at different points (e.g. mobile equipment, network switches, etc.). It is usually defined as the time between sending and reception of certain protocol messages. However, the exact message types and even protocols to be used will vary with the specific context.

More specifically, when applying abstract QoS parameter definitions to real contexts of service usage, reference shall be made to specific **reference points** and, at these reference points, to specific **protocol layers** or **service access points**, where the respective quantities shall be observed in that context.

The combination of a particular reference point with a particular protocol layer/service access point at that reference point is referred to as "point of observation (PO)" in the present document.

NOTE 1: In general, the interfaces between two network elements or between two functional entities connected via a physical layer are commonly referred to as reference points.

NOTE 2: This concept is motivated by and closely related to the concept of a *point of control and observation (PCO)* as defined in Recommendation ITU-T X.290 [i.2]. In the context of the present document, however, it is not required that sending of a message or service primitive is performed **actively** by the measurement system at the point of observation (although this might be preferable for modelling user behaviour). For the purpose of defining a specific event realization in terms of a protocol message/primitive it does not matter whether this message/primitive is actively triggered or just observed. The term "control" is therefore omitted in the present document.

For the same (abstractly characterized) measurable quantity several different points of observation - and thus several different final definitions may be specified, mainly motivated by the following reasons:

- To deal with different types of access networks, like GSM, UMTS, ISDN, etc. These differences are mainly reflected in different reference points.
- To overcome accessibility restrictions. For example, it might be easier to trace and track protocol messages in an end system at a higher protocol level than to access a bridging protocol at a network element located in a closed network of an operator. As another example, the air interface might not be accessible with particular measurement equipment whereas the AT interface is. Further issues might be caused by encryption, by the complexity of used protocols (e.g. binary coding of messages) or by the efforts to get knowledge of messages (e.g. direct air interface tracing).
- To extend the applicability of QoS parameter definitions to suitable network interfaces for special purposes. E.g. network operators with full access to all protocol levels at their network elements might make use of these interfaces to assess QoS aspects from a more network-centric viewpoint (see ETSI TS 102 250-7 [2] for details).

From a formal point of view this means that only the combination of quantity name and PO name unambiguously identifies a final definition. Consequently, measurement reports shall always indicate the PO where any particular quantity has been measured.

QoS parameters may depend on multiple quantities, each of which may in turn have been specified for different POs. However, it might not make sense to combine those POs arbitrarily. Appropriate restrictions should be specified as part of the QoS parameter definition in these cases.

Results from measurements of a particular QoS parameter based on different points of observation for the same abstract quantity shall not be compared directly (e.g. in benchmarking campaigns).

For the selection of appropriate points of observation during measurements, the following general recommendations apply:

- The PO should be chosen as close as possible to the user perception.
- As far as applicable, the same PO should be chosen for all quantities the QoS parameter to be measured depends on.
- Points of observation chosen for the different phases of a service usage transaction should be the same or at least equivalent with respect to user perception.

It is up to the reader to choose the points of observation suiting best the actual needs and/or situation, but any deviations from the recommendations given above should be explained.

Some frequently used points of observation are specified in clause 7 of the present document. There, the following notation is used:

<protocol or service access point acronym>@<reference point acronym>

e.g. "AT@R".

Other specifications may define additional points of observation, as appropriate. The same naming convention should be used.

4.3 Events

4.3.1 Introduction

A large class of QoS parameters can be expressed in terms of events. Prominent representatives are QoS parameters associated with the success (or failure) of a particular transaction (e.g. the setup phase of a telephony call) and QoS parameters measuring the duration of a successful transaction.

To alleviate the applicability of QoS parameter definitions to different contexts (e.g. involving different access technologies or service access points), two different, yet related, concepts of the term "event" are employed in the present document: **abstract events** and **technical events**.

4.3.2 Abstract events

An **abstract event** is an abstractly formulated proposition, which, for any given instance of a service usage (i.e. for any statistical sample) can be said to be either true or false. Here, "abstractly formulated" means that the proposition should usually not refer to technical details like particular protocol messages but express a property of the statistical sample in a generic way.

Examples, taken from telephony call setup, are:

EXAMPLE 1: *"Complete address information has been sent by the originating party".*

EXAMPLE 2: *"The CLI received by the terminating party is valid with respect to the originating party".*

Apart from the elementary question whether a particular event is true (has occurred) for a given instance of service usage, the dynamic aspect of **when** the event occurred (became true) is often of interest. More precisely, many QoS parameter definitions involve the time elapsed between the occurrences of two specified events. In the telephony examples given above, this dynamic aspect is important for the first one which marks the begin of the call setup phase, while the second example deals with the validity of the CLI and thus a "time of occurrence" is not relevant in this case.

4.3.3 Technical events

A **technical event**, in the sense of the present document, usually represents the observation that a particular protocol message or service primitive or some other well-defined piece of data has been either sent or received at a particular point of observation in the transmission system. Other kinds of technical events are sometimes relevant as well, e.g. timeouts or application-level software events. For technical events it is less common to assign a *truth* value. Instead, one speaks of the **occurrence** of such events.

4.3.4 Representations of an event at different points of observation

The concepts of **abstract** and **technical** events are related by the following interpretation: Abstract events are understood to be represented (or "realized") by technical events at particular points of observation (see clause 4.2) in the following sense: The abstract event is said to become true (at the selected point of observation) if and when the corresponding technical event occurs. Considering the examples given in clause 4.3.2, possible representations by technical events are:

EXAMPLE 1: *"ATD command with MSISDN of terminating party sent (at originating party reference point R, AT layer).".*

EXAMPLE 2: *"CC SETUP with valid CgPN received (at terminating party, Uu interface, layer 3 (NAS)).".*

Following this interpretation, different locations in a transmission system may be used to detect the occurrence of an abstract event. This is illustrated by figure 4.1 which shows the propagation of some imaginary event through the protocol stacks of a 3G mobile network.

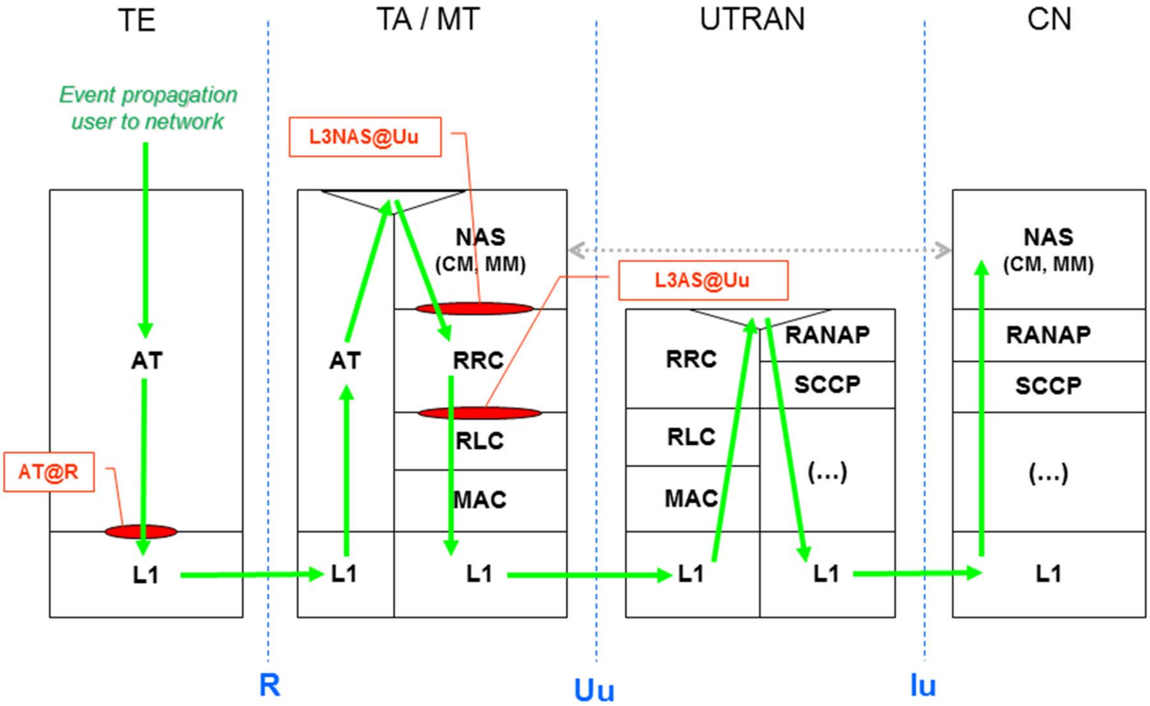

Figure 4.1: Illustration of the propagation of an event through different points of observation

The figure also visualizes the two dimensions involved in the specification of a PO:

- the horizontal position on the "user-to-network" axis, given by the reference point;
- the vertical position in the stack of the protocol layers involved at the user side of the reference point.

In fact, the logical propagation of an abstract event is not always realized by a single transmission (as the illustration might suggest). Several consecutive message flows may be involved. This shall be considered when determining which technical event, at a given reference point, provides the best approximation to the user experience.

EXAMPLE 3: Consider the abstract event given in example 1 of clause 4.3.2 and its technical realization at the UMTS Uu interface: Here, the user pressing the "connect" button usually does not directly trigger a "CC SETUP" NAS message but a radio bearer establishment procedure first. Therefore, "RRC CONNECTION REQUEST sent" is usually taken as the technical representation of the abstract event, although it is on a lower level in the protocol stack and in fact does not carry "complete address information".

Any abstract event specified according to the present document should be defined from a full end-to-end perspective, motivated by a particular user interaction or perception, respectively. This user perspective should be indicated together with the event definition. Nevertheless, user perspective and abstract event should not be mistaken to be the same - the respective user interaction should rather be taken as one possible realization of the abstract event, namely at the interface between man and machine.

4.3.5 Two-stage specification of QoS parameters defined in terms of events

Using the differentiated concept of events described in clauses 4.3.2 to 4.3.4, QoS parameters which can be expressed in terms of events, shall be defined in two stages:

1) The generic definition of the QoS parameter shall be formulated in terms of abstract events. At this stage, the definition shall not refer to the details of a particular technical context.

2) When applying the generic definition to a particular context (e.g. involving a particular access technology or a specific service access point), a PO (see clause 7) shall be selected for each abstract event involved, and the abstract event shall be formally replaced by the technical event that has been specified to represent the abstract event at the respective PO.

The separation into two stages helps to isolate the generic aspects of a QoS parameter definition and its validity from the technical details of its application to particular access contexts. This in turn alleviates giving harmonized definitions of QoS parameters whose applicability can easily be extended to additional contexts/scenarios whenever the need arises.

4.4 Transactions and their trigger events

Any action/procedure/process with a well-defined start and a well-defined end may be modelled as a transaction T, characterized by two specified events $E_{\mathrm{start},T}$ and $E_{\mathrm{end},T}$, which define its start and end, respectively. This kind of abstraction corresponds to a point of view where one is not interested in the details of a particular process, but only in its (intended) result, represented by the event $E_{\mathrm{end},T}$.

NOTE 1: For practical reasons, the end event $E_{\mathrm{end},T}$ of a transaction T is said to be true only if it has occurred within a predefined timeout period after the start of the transaction.

$E_{\mathrm{start},T}$ is commonly called the **start trigger** of the transaction T, and $E_{\mathrm{end},T}$ is called the **end trigger** of the transaction T. More precisely, when a given event A is used to specify the start of a transaction T, it is said to play the role of the start trigger for this transaction. The same holds for the end trigger role, correspondingly.

NOTE 2: The same event A can play the role of a start (or end) trigger for different transactions, and the same event A can play the role of a start trigger for one transaction, and the role of an end trigger for a different transaction.

A transaction is said to be **successful** if a specified **success condition** $E_{\mathrm{success},T}$ is true after its completion. Of course, $E_{\mathrm{success},T}$ always implies $E_{\mathrm{end},T}$. In many cases, the success condition will in fact be identical to the end trigger. There are cases, however, where the event that marks the end of a transaction on the timeline needs to be distinguished from the success condition. For example, asserting the success may require additional criteria to be checked after the end trigger has already occurred.

To summarize, a transaction T is said to be successful **if** $E_{\mathrm{success},T}$ is true and **when** $E_{\mathrm{end},T}$ occurs (within a predefined timeout period after the start of the transaction).

NOTE 3: As a consequence of note 1, the success condition $E_{\mathrm{success},T}$ is said to be false if the end trigger $E_{\mathrm{end},T}$ has not occurred within a predefined timeout period after the start of the transaction.

Whenever a transaction T has been started (i.e. its start event $E_{\mathrm{start},T}$ has become true) but does not end with the intended result within a predefined timeout period, the transaction is said to have **failed**.

NOTE 4: This definition does not distinguish between "explicit" failures, which may be detected technically, e.g. by the occurrence of some protocol message indicating a failure, and those cases where the intended end event has not been reached within a predefined timeout period.

5 Symbols and formulas for common types of QoS parameters

5.1 Introduction

Many QoS parameters are of the same type (e.g. ratios, time durations, etc.) which means that their defining equations are structurally identical. Likewise, many QoS parameter definitions refer to basic measurable quantities of the same type (e.g. event indicators, times, etc.). The following clauses introduce frequently used types of basic quantities and QoS parameters and specify the associated generic symbols and calculation formulas. These symbols shall be used, and the corresponding formulas shall be referred to or copied verbatim, in other sub-parts of ETSI TS 102 250-8 [i.3] or any other document defining individual QoS parameters of the respective type according to the present document.

5.2 Basic measurable quantities associated with events

5.2.1 Event indicators

An event indicator has the value 1 if the corresponding event has occurred and the value 0 otherwise.

Formally, to any event A, an elementary measurable quantity I_A is assigned. It is called **indicator** of the event A:

$$I_A := \begin{cases} 1 & \text{if } A \text{ is true / has occurred} \\ 0 & \text{else} \end{cases} \tag{5.1}$$

5.2.2 Event times

An event time is the point in time when the corresponding event has occurred.

For any event A, the time of occurrence of the event is denoted by t_A.

This time is measured relative to some fixed, predefined origin of time. As only time **differences** occur in the definition of QoS parameters, an arbitrary origin may be chosen, as long as the same origin is used for all event times in a particular measurement cycle.

If the event A has not occurred in a measurement cycle, t_A is not defined for that cycle.

5.3 Generic aggregation formulas for sample averages

5.3.1 Introduction

The following clauses specify generic symbols and formulas for the most common type of aggregation, usually called "sample averaging". Throughout the entire clause 5.3 it is assumed that a series of $N > 0$ measurement cycles has been performed. All the aggregations introduced in the following clauses depend on this number N, which can be seen on the right-hand sides of their definitions. For simplicity, however, this dependency is not noted explicitly in the symbols representing the aggregation entities.

5.3.2 Event count

The event count for an event *A* in a series of *N* measurement cycles is defined as the number of measurement cycles where *A* has occurred. It is denoted by $n(A)$.

5.3.3 Sample average

In the present document, the sample average is given by the arithmetic mean.

Formally, the sample average of a measurable quantity *Q* in a series of *N* measurement cycles is denoted by $\bar{Q}$ and defined by the arithmetic mean:

$$\bar{Q} := \frac{1}{N} \sum_{i=1}^{N} q_i \qquad (5.2)$$

where:

q_i is the value of *Q* measured in the *i*-th cycle.

NOTE: Formula (5.2) is also applicable if *Q* is itself a function of other measurable quantities, in particular for sums or products of such quantities.

5.3.4 Conditional sample average

A conditional sample average is an average over a subset of the whole series of measurements, where the subset is selected according to a given condition.

Formally, in a series of *N* measurement cycles, the sample average of a measurable quantity *Q* under the condition that an event *B* has occurred (or short: "under the condition B") is defined and denoted as:

$$\overline{Q|B} := \frac{1}{n(B)} \sum_{\{i:\, B \text{ is true}\}} q_i \qquad (5.3)$$

where:

q_i is the value of *Q* measured in the *i*-th cycle;

$n(B)$ is the event count for event *B*, see clause 5.3.2;

and the summation on the right-hand side is restricted to those of the *N* cycles where *B* is true.

The conditional sample average is defined only if *B* is true for at least one of the *N* cycles.

5.4 Ratios

5.4.1 Sample ratio of an event

A sample ratio of an event is the number of its occurrences divided by the total number of cycles.

Formally, a sample ratio is the relative frequency of occurrence of some specified event *A* in a series of *N* measurement cycles and is denoted as $r(A)$. It is usually taken as an estimate of the probability of occurrence of the event *A*, or, in other words, of the probability of the proposition expressed by the abstract event.

A ratio cannot be determined by performing just a single measurement. The corresponding measurable quantity is the indicator of the event *A*.

$$r(A) := \frac{n(A)}{N} \qquad (5.4)$$

where:

$n(A)$ is the event count for event *A*, see clause 5.3.2.

5.4.2 Conditional ratio of an event

A conditional ratio is a sample ratio over a subset of the whole series of measurements, where the subset is selected according to a given condition.

Formally, a conditional ratio is the relative frequency of occurrence, in a series of *N* measurement cycles, of an event *A* under the condition that another event *B* has occurred/is true. It is usually taken as an estimate of the corresponding conditional probability. The conditional ratio is denoted by $r(A|B)$:

$$r(A|B) := \frac{r(A \wedge B)}{r(B)} = \frac{n(A \wedge B)}{n(B)} \quad (5.5)$$

where:

$A \wedge B$ denotes the logical AND combination of the events *A* and *B;*

$r(A \wedge B)$ and $r(B)$ are sample ratios of occurence as defined by equation (5.4);

$n(A \wedge B)$ and $n(B)$ are event counts, see clause 5.3.2.

5.4.3 Success ratio of a transaction

A success ratio is the number of successful transactions divided by the total number of started transactions.

Formally, the success ratio of a transaction T (see clause 4.4), in a series of *N* measurement cycles, is denoted by $r_{\text{success},T}$ and given by the conditional ratio of its success event $E_{\text{success},T}$ with respect to its start event $E_{\text{start},T}$. The success condition of a transaction can only become true if the transaction has been started at all, so that, formally, $E_{\text{start},T} \wedge E_{\text{success},T} = E_{\text{success},T}$. Thus, the general formula for conditional ratios, equation (5.5), in this case simplifies to:

$$r_{\text{success},T} := \frac{n(E_{\text{success},T})}{n(E_{\text{start},T})} \quad (5.6)$$

where:

$E_{\text{start},T}$ and $E_{\text{success},T}$ denote the events defining the start and success of the transaction *T*, respectively;

$n(E_{\text{start},T})$ and $n(E_{\text{success},T})$ are the corresponding event counts, see clause 5.3.2.

5.4.4 Failure ratio of a transaction

A failure ratio is the complement of the corresponding success ratio.

Formally, the failure ratio of a transaction *T*, in a series of *N* measurement cycles, is denoted as $r_{\text{fail},T}$ and defined by:

$$r_{\text{fail},T} := 1 - r_{\text{success},T} \quad (5.7)$$

where:

$r_{\text{success},T}$ is the success ratio of the transaction *T*, see equation (5.6).

NOTE 1: For most non-trivial cases it is hard - if possible at all - to give a closed characterization of a "failure" event for a given transaction *T* (see clause 4.4). This would usually require knowledge of all possible failure scenarios and how these can be detected technically. Furthermore, the number and kind of possible failures may change over time and vary with different network implementations. For these reasons, no attempt is made to give a direct definition of a failure ratio in terms of failure events, using equation (5.5).

NOTE 2: This formula assumes that any "inconclusive" cases, where the transaction *T* did neither succeed nor fail with a reason attributable to the system under test, have been excluded from the measurement data (i.e. are not part of the *N* cycles under consideration), e.g. by appropriate post-processing methods.

5.5 Durations

5.5.1 Duration of a transaction

A duration is the time elapsed between two events.

Formally, the duration of a transaction *T* is denoted by Δt_T and defined as the time elapsed between the two specified events $E_{\text{start},T}$ and $E_{\text{end},T}$ which define the start and (intended) end, respectively, of the transaction:

$$\Delta t_T := t_{E_{\text{end},T}} - t_{E_{\text{start},T}} \tag{5.8}$$

NOTE 1: The reference to the individual event times is conceptual. It does not imply that measuring the individual times is mandatory. Measuring a duration could also be performed, e.g., by starting a timer when the start event $E_{\text{start},T}$ occurs, stopping the timer when the end event $E_{\text{end},T}$ occurs, and reading out the elapsed time directly.

NOTE 2: A duration is only defined if both involved events have actually occurred in the respective measurement cycle.

5.5.2 Mean duration of a transaction

A mean duration is the average of the measured durations, calculated over the subset of successful transactions.

Formally, the mean duration of a transaction *T*, in a series of *N* measurement cycles, is denoted by $\widetilde{\Delta t_T}$ and given by the conditional sample average of the duration Δt_T with respect to the success condition $E_{\text{success},T}$ of the transaction. In other words, when calculating the mean duration $\widetilde{\Delta t_T}$, only those of the *N* cycles are taken into account for which the transaction *T* was started and completed successfully within a predefined timeout period:

$$\widetilde{\Delta t_T} := \overline{\Delta t_T | E_{\text{success},T}} = \frac{1}{n(E_{\text{success},T})} \sum_{\{i:\, T \text{ successful}\}} \Delta t_{T,i} \tag{5.9}$$

where:

Δt_T is the duration of the transaction *T* as defined by equation (5.8);

$E_{\text{success},T}$ is the success condition of the transaction *T;*

$\overline{\bullet | \bullet}$ denotes conditional sample averaging, see equation (5.3);

$n(E_{\text{success},T})$ is the number of successful transactions *T;*

$\Delta t_{T,i}$ is the value of Δt_T measured in the *i*-th cycle;

and the summation on the right-hand side is restricted to those of the *N* cycles where the transaction *T* was successful.

5.6 Data amount and data rates

5.6.1 Data amount transferred in a transaction

A data amount is understood as the amount of transferred data.

Formally, the data amount transferred in a transaction *T* is denoted by ΔD_T.

5.6.2 Data rate of a transaction

A data rate is the amount of data transferred in a transaction, divided by the duration of this transaction.

Formally, the data rate of a data transfer transaction *T*, is denoted by d_T and defined by:

$$d_T := \frac{\Delta D_T}{\Delta t_T} \tag{5.10}$$

where:

ΔD_T denotes the amount of data transferred in the transaction *T;*

Δt_T is the duration of the transaction *T*, see clause 5.5.1.

NOTE: The instantaneous rate at which data is transferred is a highly fluctuating quantity. The definition of the data rate d_T involves an implicit aggregation over the duration Δt_T. This type of aggregation should not be confused with a sample average over several transactions.

5.6.3 Mean data rate of a transaction

A mean data rate is the average of the measured data rates, calculated over the subset of successful transactions.

Formally, the mean data rate of a transaction *T*, in a series of *N* measurement cycles, is denoted by $\widetilde{d_T}$ and given by the conditional sample average of the data rate d_T with respect to the success condition $E_{\text{success},T}$ of the transaction. In other words, when calculating the mean data rate $\widetilde{d_T}$, only those of the *N* cycles are taken into account for which the transaction *T* was started and completed successfully within a predefined timeout period:

$$\widetilde{d_T} \coloneqq \overline{d_T | E_{\text{success},T}} = \frac{1}{n(E_{\text{success},T})} \sum_{\{i:\,T \text{ successful}\}} d_{T,i} \tag{5.11}$$

where:

d_T is the data rate of the transaction *T* as defined by equation (5.10);

$E_{\text{success},T}$ is the success condition of the transaction *T;*

$\overline{\bullet | \bullet}$ denotes conditional sample averaging, see equation (5.3);

$n(E_{\text{success},T})$ is the number of successful transactions *T;*

$d_{T,i}$ is the value of d_T measured in the *i*-th cycle;

and the summation on the right-hand side is restricted to those of the *N* cycles where the transaction *T* was successful.

6 Phases of service usage and associated QoS parameters

6.1 Segmentation into phases

In many cases, the usage of a service may be looked upon as being represented by a main transaction (e.g. a telephone call or an MMS transmission) which can be broken down to a sequence of several successive sub-transactions, each of which is relevant for the user's perception of the service. These sub-transactions are called **phases** (of the usage of the respective service) in the present document.

EXAMPLE: A telephony call is usually divided into a call setup phase, a call phase, and a call termination phase.

It is emphasized that not every technically conceivable (sub-)transaction related to the usage of a service shall be treated as a phase in the sense described in the present clause. The essential criterion is that the sub-transaction is perceived by the end user, and that its success, and possibly timing, is directly affecting the user's experience of the service under consideration.

Being modelled as transactions, the phases of a service usage shall be defined by specifying their start and end trigger events according to clause 4.4.

A *seamless* segmentation of a service transaction into phases consists of a sequence of phases which:

- do not overlap in time; and
- do not leave any gap between successive phases.

In other words, for any two successive phases T_1 and T_2 in a seamless segmentation, the end trigger of the first phase is given by the same event as the start trigger of the subsequent phase: $E_{\mathrm{end},T_1} = E_{\mathrm{start},T_2}$.

When defining QoS parameters for a particular service according to the present document, the service usage should be split into meaningful phases. Whether this segmentation follows the seamless pattern just described or not should be indicated in the respective service clause.

QoS parameters for a particular phase should be defined together in a clause corresponding to that phase.

QoS parameters which cannot be attributed unambiguously to a particular phase should be defined in a separate clause.

6.2 Success of a phase

The QoS parameter definition for the success of a specified phase named *T* is given as follows:

Measurable quantity: the **success indicator** $I_{E_{\mathrm{success},T}}$ of the phase, see clause 5.2.1:

$$I_{E_{\mathrm{success},T}} := \begin{cases} 1 & \text{if } E_{\mathrm{success},T} \text{ is true} \\ 0 & \text{else} \end{cases} \tag{6.1}$$

Aggregated QoS parameter: the **success ratio** $r_{\mathrm{success},T}$ of the phase, see clause 5.4.3:

$$r_{\mathrm{success},T} := \frac{n(E_{\mathrm{success},T})}{n(E_{\mathrm{start},T})} \tag{6.2}$$

The corresponding failure ratio is implicitly defined as well, see clause 5.4.4.

In these equations, $E_{\mathrm{start},T}$ and $E_{\mathrm{success},T}$ shall be substituted by the events specified for start trigger and success condition, respectively, in the definition of the phase. Substitution with the abstract events yields the generic definition (stage 1). Replacement of the abstract events by the corresponding technical events, as defined in the event specification for a specific context, yields the final definition (stage 2) for that context.

6.3 Duration of a phase

For the success of a phase named *T*, defined by the events $E_{\mathrm{start},T}$ ("start trigger"), $E_{\mathrm{end},T}$ ("end trigger") and $E_{\mathrm{success},T}$ (success condition, which may be identical to $E_{\mathrm{end},T}$), the following QoS parameter definition applies:

The QoS parameter definition for the duration of a specified phase named *T* is given as follows:

Measurable quantity: the **duration** Δt_T of the phase, see clause 5.5.1:

$$\Delta t_T := t_{E_{\mathrm{end},T}} - t_{E_{\mathrm{start},T}} \tag{6.3}$$

Aggregated QoS parameter: the **mean duration** $\widetilde{\Delta t_T}$ of the phase, see clause 5.5.2:

$$\widetilde{\Delta t_T} \coloneqq \overline{\Delta t_T | E_{\mathrm{success},T}} = \frac{1}{n(E_{\mathrm{success},T})} \sum_{\{i:\, T \text{ successful}\}} \Delta t_{T,i} \tag{6.4}$$

In these equations, $E_{\mathrm{start},T}$, $E_{\mathrm{end},T}$ and $E_{\mathrm{success},T}$ shall be substituted by the events specified for start trigger, end trigger and success condition, respectively, in the definition of the phase. Substitution with the abstract events yields the generic definition (stage 1). Replacement of the abstract events by the corresponding technical events, as defined in the event specification for a specific context, yields the final definition (stage 2) for that context.

6.4 Other QoS parameters related to a phase

Depending on the particular service and the particular phase, possibly other QoS parameters related to that phase may be defined, whose calculation is not necessarily based on the start and end triggers. For those QoS parameters, the start and end triggers usually will provide the time frame for carrying out the corresponding base measurements.

7 Points of observation

7.1 Overview

7.1.1 GERAN/UTRAN/E-UTRAN network access model

With respect to user-perceived QoS, many interactions refer to the basic model shown in figure 7.1. It shows a typical mobile device (MS in GSM or UE in UMTS and LTE) and its internal structure. For further information see ETSI TS 124 002 [3].

Different interfaces are identified here:

- Many human interactions with a mobile device are executed via the keypad, the display, the acoustic elements like microphone and loudspeaker and some switches at the case of the device. All these elements form the man-machine interface (MMI).

- In terms of QoS and especially active probing, human interactions are often simulated for automation purposes. In this case, an application represents the human user and invokes interactions with the mobile termination (MT) part of the mobile device via the terminal equipment (TE) and the terminal adapter (TA). The collaboration of TE and MT represents classical network termination functionality like it is used in ISDN technology.

- The interworking between the SIM or U-SIM and the MT is defined via the SIM-ME interface in GSM and the Cu interface in UMTS/LTE, respectively.

- Finally, on the network side, MT interacts with the wireless network via the air interface (Um in GSM and Uu in UMTS and LTE).

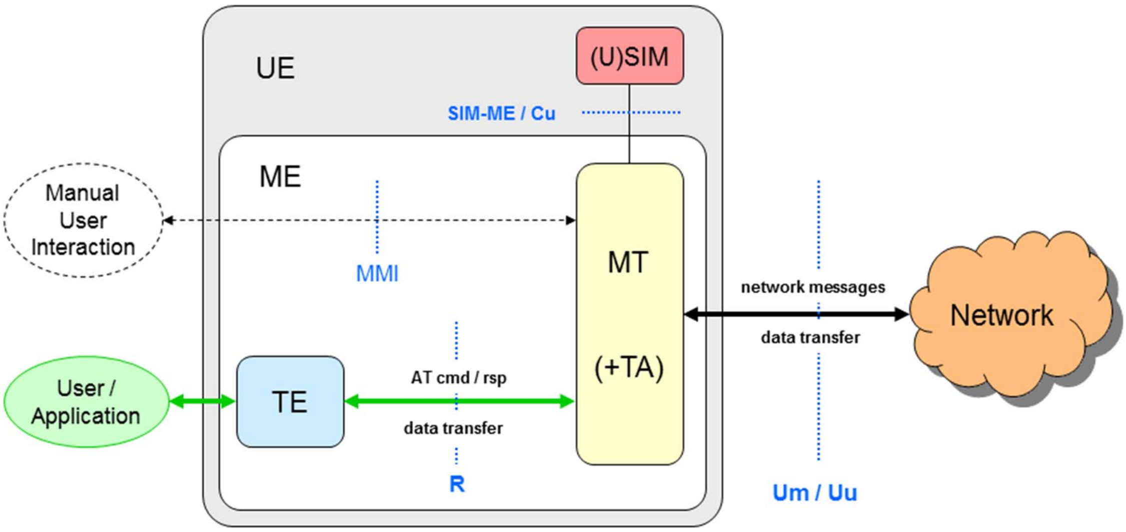

Figure 7.1: Standard 2G/3G/4G network access model

The conceptual approach to differentiate between the TE and the TA/MT introduces the reference point R. This reference point allows the exchange of AT commands and AT responses and acts as a relay for packet switched data. The AT command set is used to control the MT, to retrieve information about the MS or UE and to control the interaction with the mobile network.

At this stage it is important to notice that only the combination of a certain protocol (AT command set here) and a certain location within the architecture (R reference point in this case) represents a unique point of observation.

7.1.2 2G CS domain

Figure 7.2 illustrates the points of observation defined in the present document for the 2G CS domain. For further information see ETSI TS 143 051 [4].

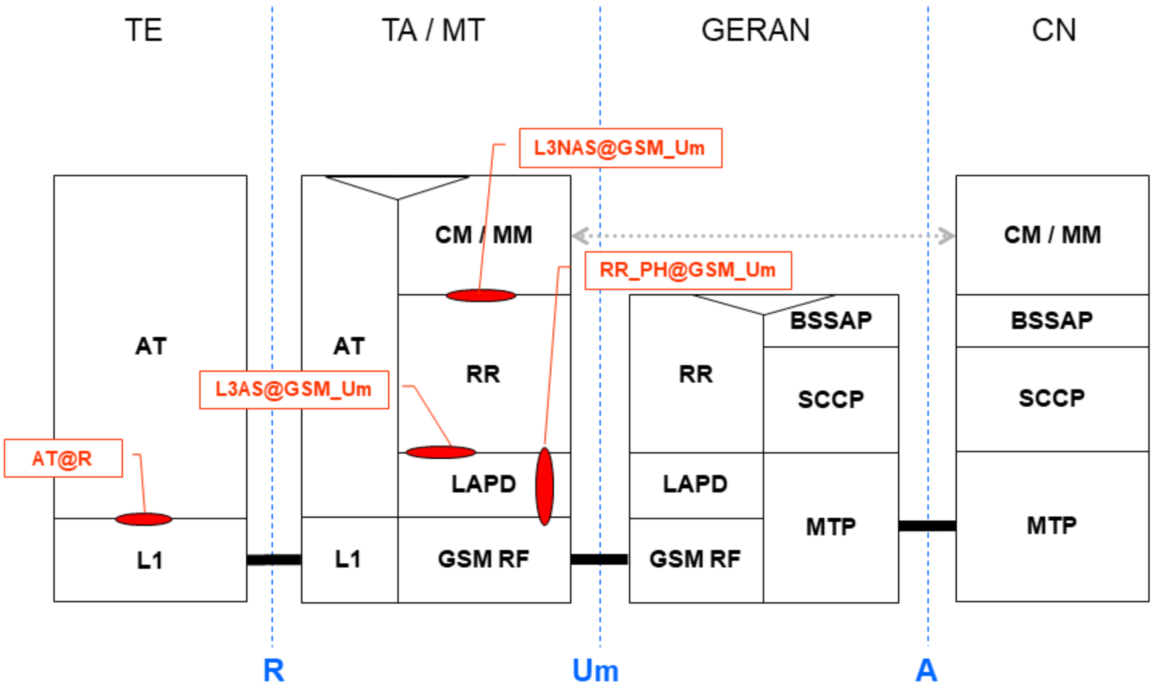

Figure 7.2: 2G CS domain

7.1.3 2G PS control plane

Figure 7.3 illustrates the points of observation defined in the present document for the 2G PS control plane. For further information see ETSI TS 123 060 [5].

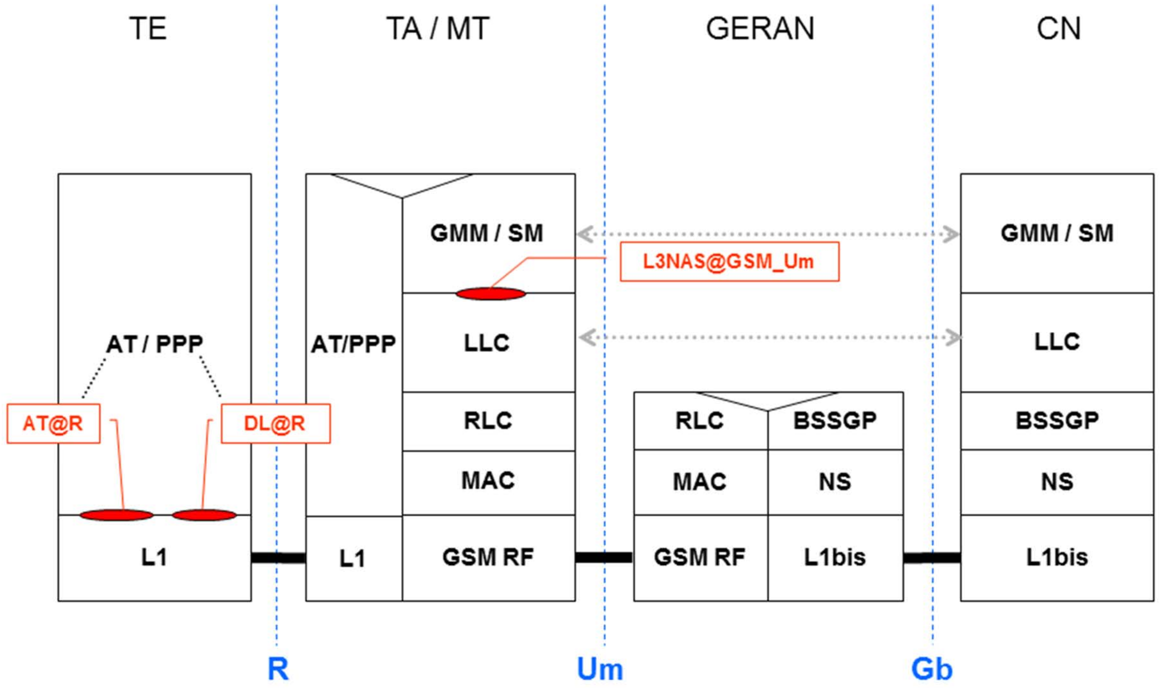

Figure 7.3: 2G PS control plane

7.1.4 2G PS user plane

Figure 7.4 illustrates the points of observation defined in the present document for the 2G PS user plane. For further information see ETSI TS 123 060 [5].

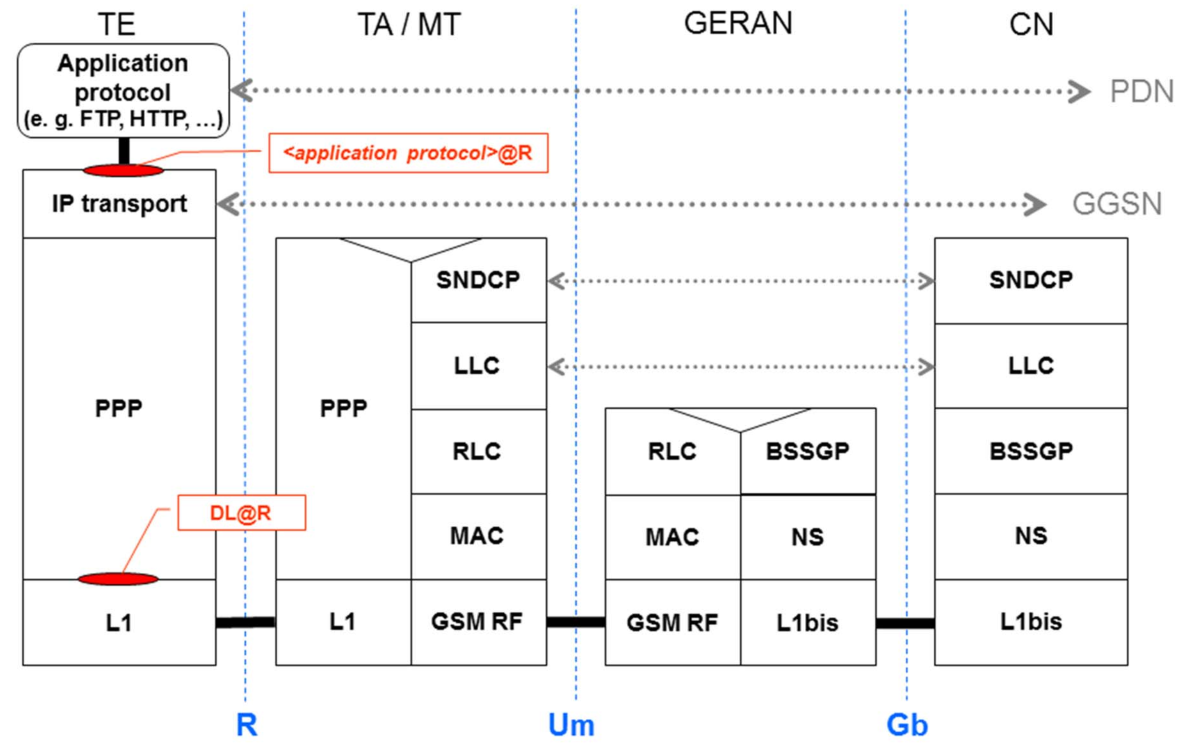

Figure 7.4: 2G PS user plane

7.1.5 3G CS domain and PS control plane

Figure 7.5 illustrates the points of observation defined in the present document for the 3G CS domain and the PS control plane. For further information see ETSI TS 123 060 [5].

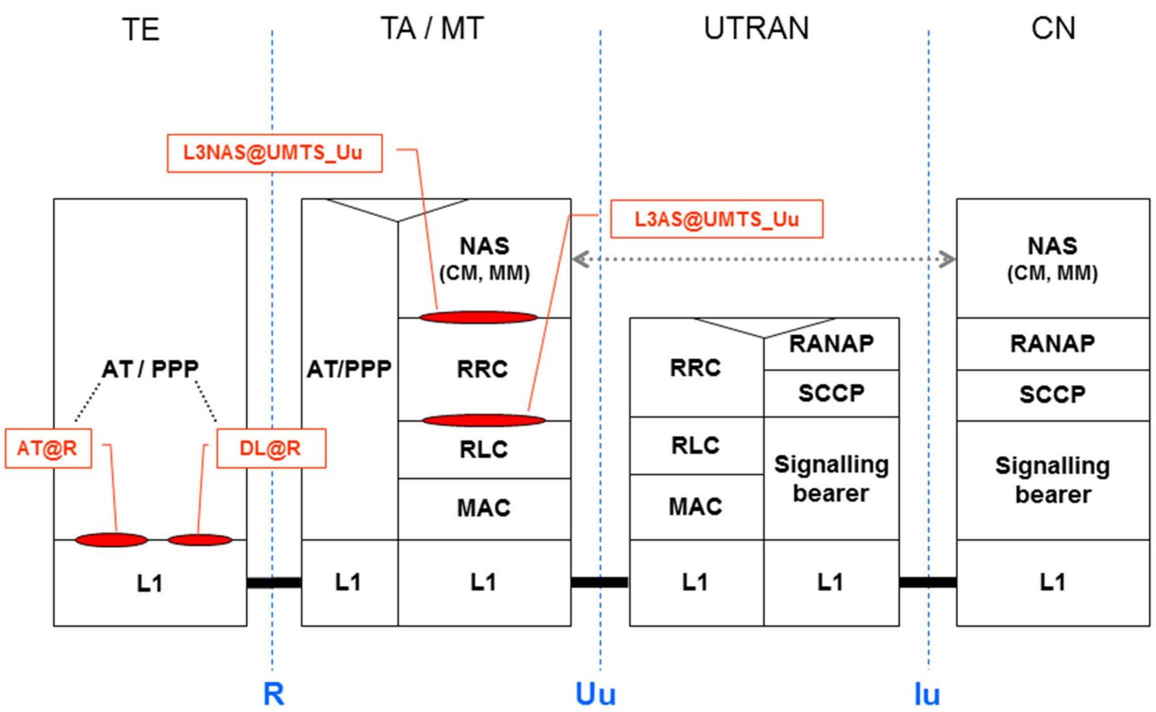

Figure 7.5: 3G CS domain and PS control plane

7.1.6 3G PS user plane

Figure 7.6 illustrates the points of observation defined in the present document for the 3G PS user plane. For further information see ETSI TS 123 060 [5].

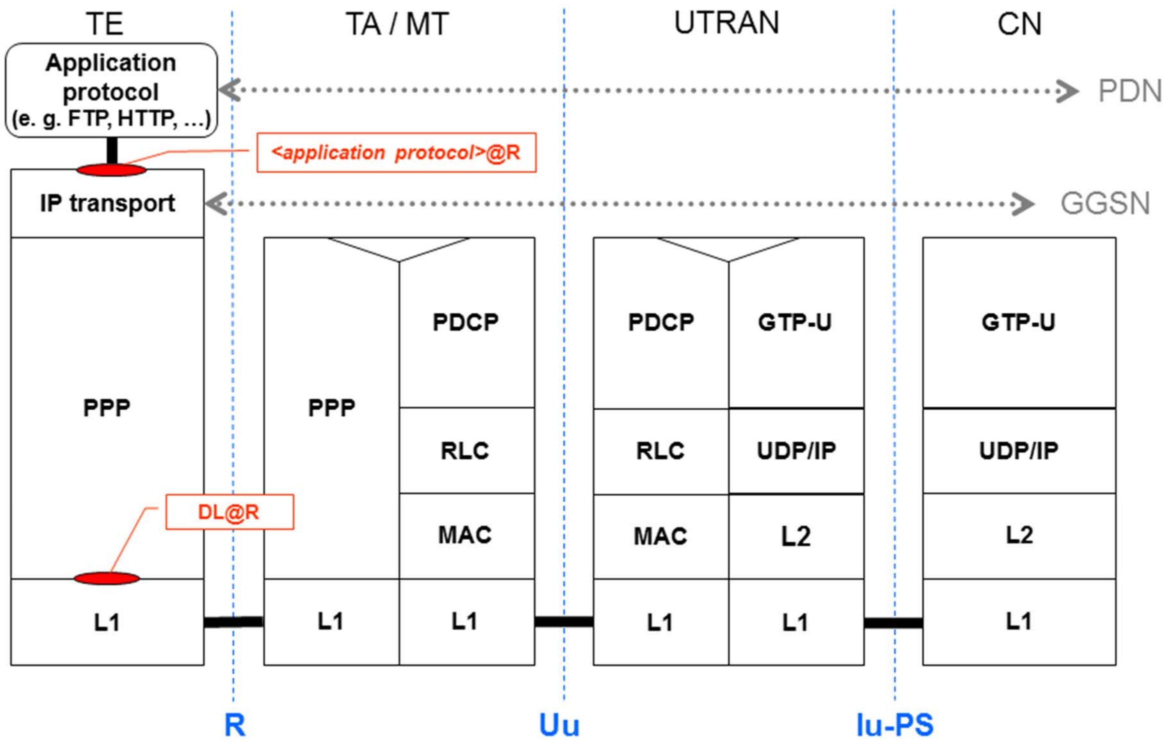

Figure 7.6: 3G PS user plane

7.1.7 4G PS control plane

Figure 7.7 illustrates the points of observation defined in the present document for the 4G PS control plane. For further information see ETSI TS 123 401 [6].

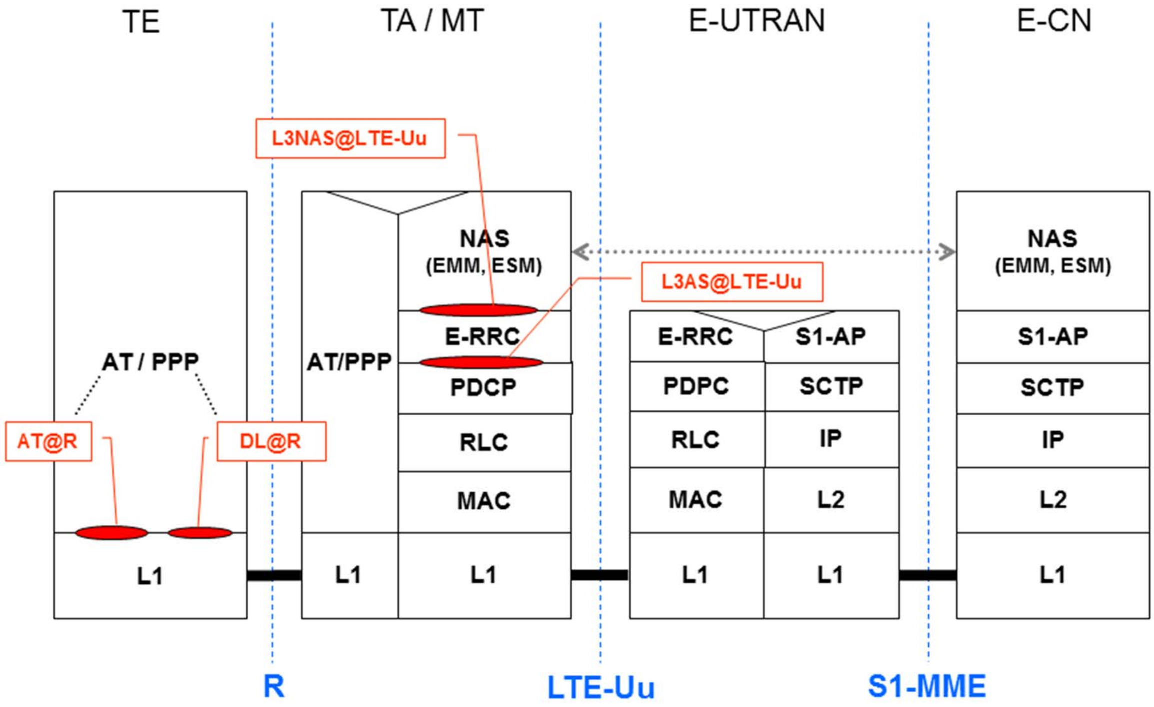

Figure 7.7: 4G PS control plane

7.1.8 4G PS user plane

Figure 7.8 illustrates the points of observation defined in the present document for the 4G PS user plane. For further information see ETSI TS 123 401 [6].

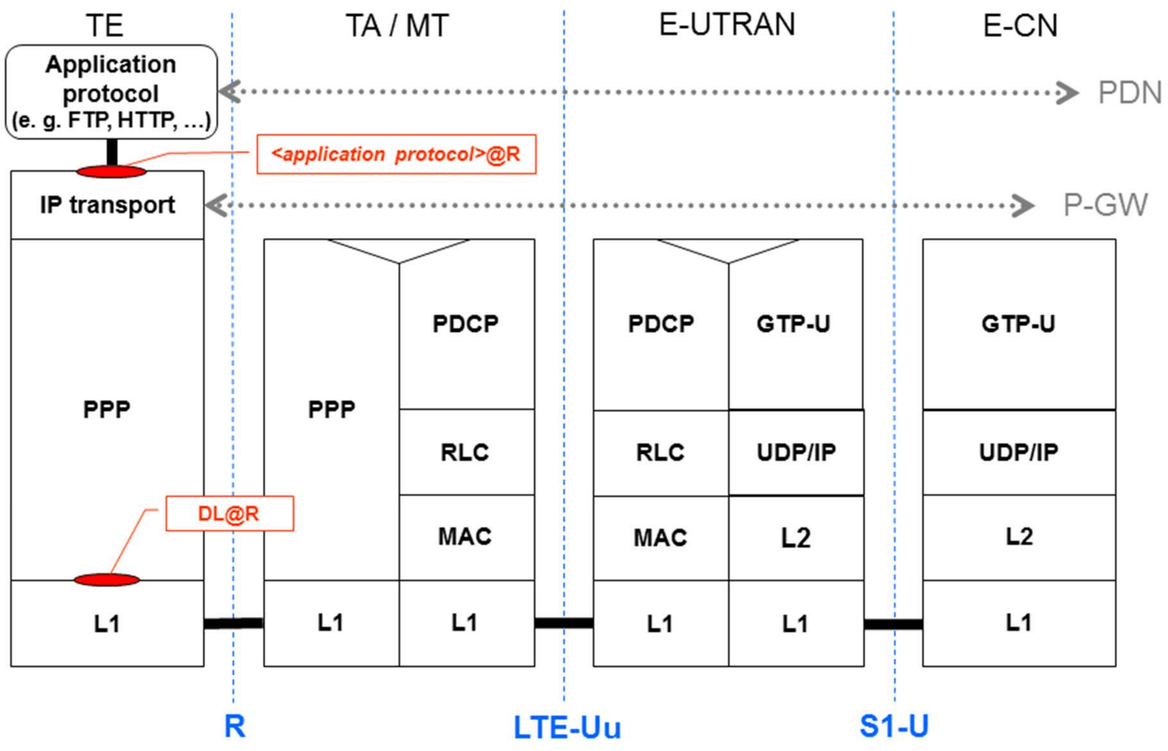

Figure 7.8: 4G PS user plane

7.2 Reference points and associated points of observation

7.2.1 GSM/UMTS/LTE reference point R

7.2.1.1 Specification of the reference point

The reference point R is specified in ETSI TS 124 002 [3].

7.2.1.2 AT@R

ID	AT@R
Description	AT command/response service access point at the R reference point.
Related overview figures	For GSM, see figure 7.2 and figure 7.3, for UMTS, see figure 7.5, for LTE, see figure 7.7.
Related standards	ETSI TS 124 002 [3], ETSI TS 127 007 [7].

7.2.1.3 DL@R

ID	DL@R
Description	Service access point between datalink (e.g. PPP) and physical layer (L1) at the R reference point.
Related overview figures	For GSM, see figure 7.3 and figure 7.4, for UMTS, see figure 7.5 and figure 7.6, for LTE, see figure 7.7 and figure 7.8.
Related standards	ETSI TS 127 060 [8].
NOTE: This is the access point used e.g. by packet capture libraries.	

7.2.1.4 IMAP4@R

ID	IMAP4@R
Description	IMAP4 protocol layer at the R reference point.
Related overview figure	For GSM, see figure 7.4, for UMTS, see figure 7.6, for LTE, see figure 7.8.
Related standards	IETF RFC 3501 [9].
NOTE: This is e.g. the access point of some sockets API.	

7.2.1.5 POP3@R

ID	POP3@R
Description	POP3 protocol layer at the R reference point.
Related overview figure	For GSM, see figure 7.4, for UMTS, see figure 7.6, for LTE, see figure 7.8.
Related standards	IETF RFC 1939 [10].
NOTE: This is e.g. the access point of some sockets API.	

7.2.1.6 SMTP@R

ID	SMTP@R
Description	SMTP protocol layer at the R reference point.
Related overview figure	For GSM, see figure 7.4, for UMTS, see figure 7.6, for LTE, see figure 7.8.
Related standards	IETF RFC 5321 [11].
NOTE: This is e.g. the access point of some sockets API.	

7.2.1.7 MLME@R

ID	MLME@R
Description	MLME service access point between the Station Management Entity (SME) and the Media Access Control (MAC).
Related overview figure	n.a.
Related standards	IEEE Std 802.11 [12].

7.2.1.8 MIC@R

ID	MIC@R
Description	Microphone electrical interface at the R reference point.
Related overview figure	n.a.
Related standards	n.a.

7.2.1.9 SPEAKER@R

ID	SPEAKER@R
Description	Speaker electrical interface at the R reference point.
Related overview figure	n.a.
Related standards	n.a.

7.2.1.10 CAMERA@R

ID	CAMERA@R
Description	Camera capture interface at the R reference point.
Related overview figure	n.a.
Related standards	n.a.

7.2.1.11 VIDEO@R

ID	VIDEO@R
Description	Video display interface at the R reference point.
Related overview figure	n.a.
Related standards	n.a.

7.2.1.12 DNS@R

ID	DNS@R
Description	DNS protocol layer at the R reference point.
Related overview figure	For GSM, see figure 7.4, for UMTS, see figure 7.6, for LTE, see figure 7.8.
Related standards	IETF RFC 1034 [13], IETF RFC 1035 [14].
NOTE: This is e.g. the access point of some sockets API.	

7.2.1.13 FTP@R

ID	FTP@R
Description	FTP protocol layer at the R reference point.
Related overview figure	For GSM, see figure 7.4, for UMTS, see figure 7.6, for LTE, see figure 7.8.
Related standards	IETF RFC 959 [15].
NOTE: This is e.g. the access point of some sockets API.	

7.2.1.14 SIP@R

ID	SIP@R
Description	SIP protocol layer at the R reference point.
Related overview figure	For GSM, see figure 7.4, for UMTS, see figure 7.6, for LTE, see figure 7.8.
Related standards	IETF RFC 3261 [16].
NOTE: This is e.g. the access point of some sockets API.	

7.2.1.15 RTP@R

ID	RTP@R
Description	RTP protocol layer at the R reference point.
Related overview figure	For GSM, see figure 7.4, for UMTS, see figure 7.6, for LTE, see figure 7.8.
Related standards	IETF RFC 3550 [17], IETF RFC 3711 [18].
NOTE: This is e.g. the access point of some sockets API.	

7.2.1.16 RTCP@R

ID	RTCP@R
Description	RTCP protocol layer at the R reference point.
Related overview figure	For GSM, see figure 7.4, for UMTS, see figure 7.6, for LTE, see figure 7.8.
Related standards	IETF RFC 3550 [17], IETF RFC 3711 [18].
NOTE: This is e.g. the access point of some sockets API.	

7.2.1.17 RTSP@R

ID	RTSP@R
Description	RTSP protocol layer at the R reference point.
Related overview figure	For GSM, see figure 7.4, for UMTS, see figure 7.6, for LTE, see figure 7.8.
Related standards	IETF RFC 7826 [19].
NOTE: This is e.g. the access point of some sockets API.	

7.2.1.18 HTTP@R

ID	HTTP@R
Description	HTTP protocol layer at the R reference point.
Related overview figure	For GSM, see figure 7.4, for UMTS, see figure 7.6, for LTE, see figure 7.8.
Related standards	IETF RFC 7230 [20].
NOTE: This is e.g. the access point of some sockets API.	

7.2.1.19 DHCP@R

ID	DHCP@R
Description	DHCP protocol layer at the R reference point.
Related overview figure	For GSM, see figure 7.4, for UMTS, see figure 7.6, for LTE, see figure 7.8.
Related standards	IETF RFC 2131 [21].
NOTE: This is e.g. the access point of some sockets API.	

7.2.2 TETRA reference point RT

7.2.2.1 Specification of the reference point

The reference point RT is specified in ETSI EN 300 392-5 [22].

7.2.2.2 AT@RT

ID	AT@RT
Description	AT command/response service access point at the TETRA RT reference point.
Related overview figure	n.a.
Related standards	ETSI EN 300 392-5 [22].

7.2.2.3 AUDIO@RT

ID	AUDIO@RT
Description	Audio interface at the TETRA RT reference point.
Related overview figure	n.a.
Related standards	n.a.

7.2.3 GSM interface Um

7.2.3.1 Specification of the interface

The GSM interface Um is specified in ETSI TS 144 001 [23].

7.2.3.2 L3NAS@GSM_Um

ID	L3NAS@GSM_Um
Description	Layer 3 non-access stratum sublayer at the GSM Um interface
Related overview figure	Figure 7.2 and figure 7.3
Related standards	ETSI TS 124 007 [27]

7.2.3.3 L3AS@GSM_Um

ID	L3AS@GSM_Um
Description	Layer 3 access stratum sublayer at the GSM Um interface
Related overview figure	Figure 7.2
Related standards	ETSI TS 144 005 [25]

7.2.3.4 RR_PH@GSM_Um

ID	RR_PH@GSM_Um
Description	RR-Physical service access point at the GSM Um interface
Related overview figure	Figure 7.2
Related standards	ETSI TS 144 004 [24], ETSI TS 145 008 [26]

7.2.4 TETRA interface Um

7.2.4.1 Specification of the interface

The TETRA interface Um is specified in ETSI EN 300 392-1 [28].

7.2.4.2 SNDCP@TETRA_Um

ID	SNDCP@TETRA_Um
Description	SNDCP protocol layer at the TETRA Um interface
Related overview figure	n.a.
Related standards	ETSI EN 300 392-2 [29]

7.2.4.3 CMCE@TETRA_Um

ID	CMCE@TETRA_Um
Description	CMCE protocol layer at the TETRA Um interface
Related overview figure	n.a.
Related standards	ETSI EN 300 392-1 [28]

7.2.4.4 LLC@TETRA_Um

ID	LLC@TETRA_Um
Description	LLC protocol layer at the TETRA Um interface
Related overview figure	n.a.
Related standards	ETSI EN 300 392-1 [28]

7.2.5 UMTS interface Uu

7.2.5.1 Specification of the interface

The UMTS interface Uu is specified in ETSI TS 125 301 [30].

7.2.5.2 L3NAS@UMTS_Uu

ID	L3NAS@UMTS_Uu
Description	Layer 3 non-access stratum sublayer at the UMTS Uu interface
Related overview figure	Figure 7.5
Related standards	ETSI TS 124 007 [27]

7.2.5.3 L3AS@UMTS_Uu

ID	L3AS@UMTS_Uu
Description	Layer 3 access stratum sublayer at the UMTS Uu interface
Related overview figure	Figure 7.5
Related standards	ETSI TS 125 301 [30]

7.2.6 LTE interface Uu

7.2.6.1 Specification of the interface

The LTE interface Uu is specified in ETSI TS 123 401 [6].

7.2.6.2 L3NAS@LTE_Uu

ID	L3NAS@LTE_Uu
Description	Layer 3 non-access stratum sublayer at the LTE Uu interface
Related overview figure	Figure 7.7
Related standards	ETSI TS 124 301 [31]

7.2.6.3 L3AS@LTE_Uu

ID	L3AS@LTE_Uu
Description	Layer 3 access stratum sublayer at the LTE Uu interface
Related overview figure	Figure 7.7
Related standards	ETSI TS 136 331 [32]

7.2.7 ISDN user network interface (UNI)

7.2.7.1 Specification of the interface

The ISDN user network interface UNI is specified in Recommendation ITU-T I.411 [33], Recommendation ITU-T I.412 [34].

7.2.7.2 DSS1@UNI

ID	DSS1@UNI
Description	DSS1 protocol layer at the User-Network-Interface (UNI)
Related overview figure	n.a.
Related standards	Recommendation ITU-T Q.931 [35]

Annex A (informative): Example for applying the formulas of clause 5

A.1 A fictitious service and its QoS parameters

Let "MyData" be a fictitious service for downloading data objects of various sizes over some simple, fictitious protocol. To retrieve a desired object, a MyData client needs to connect to the MyData service first. Once access to the service has been granted, the download of the requested object starts immediately.

Consider the following abstract events:

- A = "MyData object requested"
- B = "MyData welcome received"
- C = "MyData object downloaded"

which are assumed to be observable as send/receive events of certain "MyData protocol" messages.

With that, the MyData object download is divided into the following two phases, seamlessly:

- T_1 = "MyData service access", defined by: $E_{\text{start}} = A$:, $E_{\text{end}} = E_{\text{success}} = B$
- T_2 = "MyData object download",defined by: $E_{\text{start}} = B$, $E_{\text{end}} = E_{\text{success}} = C$

The following QoS parameters are assumed to be defined:

- MyData service access success/failure and time
- MyData download success/failure and time
- MyData download rate

A.2 Elementary measurement data

Consider that a campaign of N (independent) measurement cycles has been performed for the MyData service transaction. For demonstration purposes, $N = 6$ (which would be extremely small in practice). It is assumed that each cycle starts with some initial (e.g. network access) phase, for which no QoS parameters have been defined.

The elementary measurement data for each cycle consists of values for the event indicators, for the event times and for the size of the downloaded data objects. The fictitious result set is shown in table A.1.

Table A.1: Elementary MyData measurement data

Cycle	I_A	I_B	I_C	t_A / ms	t_B / ms	t_C / ms	ΔD_{T_2} / kbit
1	1	1	1	10	110	1110	100
2	1	0	0	12	-	-	-
3	0	0	0	-	-	-	-
4	1	1	0	15	215	-	-
5	1	1	1	8	308	808	75
6	1	1	1	13	213	1013	160

For simplicity, the event times are measured from the start of the respective measurement cycle.

Cycle 3 shows a case where the initial phase failed already, so that not even the start trigger of the MyData service access has been reached.

A.3 Calculated quantities per measurement cycle

From the elementary measurements listed in table A.1, the values of the QoS parameters (i.e. the values of the corresponding measurable quantities) are now calculated for each cycle:

- The values of the event indicators are already there in table A.1
- The durations are calculated according to clause 5.5.1, which means, using the phase definitions given in clause A.1: $\Delta t_{T_1} = t_B - t_A$ and $\Delta t_{T_2} = t_C - t_B$
- The data rate is calculated according to clause 5.6.2: $d_{T_2} = \frac{\Delta D_{T_2}}{\Delta t_{T_2}} = \frac{\Delta D_{T_2}}{\mathrm{t_C} - t_B}$

The results are listed in table A.2.

Table A.2: Calculated MyData QoS parameter values per measurement cycle

Cycle	Δt_{T_1} / ms	Δt_{T_2} / ms	d_{T_2} / kbit/s
1	100	1000	100
2	-	-	-
3	-	-	-
4	200	-	-
5	300	500	150
6	200	800	200

A.4 Aggregations/Sample averages

The success ratios for the phases are calculated according to clause 5.4.3, using the phase definitions given in clause A.1:

$$r_{\text{success},T_1} = \frac{n(B)}{n(A)} = \frac{4}{5} = 80\ \% \text{ and } r_{\text{success},T_2} = \frac{n(C)}{n(B)} = \frac{3}{4} = 75\ \% \quad \text{(A.1)}$$

Here, the event counts have been determined according to clause 5.3.2. For example, $n(A)$ is the number of lines in table A.1 where $I_A = 1$.

The mean durations of the phases are calculated according to clause 5.5.2, inserting the appropriate events for the success triggers:

$$\widetilde{\Delta t_{T_1}} = \frac{1}{n(B)} \sum_{\{i:\,T_1 \text{ successful}\}} \Delta t_{T_1,i} = \frac{1}{4}(100 + 200 + 300 + 200)\ \text{ms} = 200\ \text{ms} \quad \text{(A.2)}$$

and

$$\widetilde{\Delta t_{T_2}} = \frac{1}{n(C)} \sum_{\{i:\,T_2 \text{ successful}\}} \Delta t_{T_2,i} = \frac{1}{3}(1000 + 500 + 800)\ \text{ms} \approx 767\ \text{ms} \quad \text{(A.3)}$$

Likewise, clause 5.6.3 is applied to determine the mean data rate, again using $E_{\text{success}} = C$ from the definition of the MyData object download phase T_2:

$$\widetilde{d_{T_2}} = \frac{1}{n(C)} \sum_{\{i:\,T_2 \text{ successful}\}} d_{T_2,i} = \frac{1}{3}(100 + 150 + 200)\ \text{kbit/s} = 150\ \text{kbit/s} \quad \text{(A.4)}$$

Annex B (informative): Guidelines and templates for specification

B.1 Document structure

This annex outlines a possible document structure for the formalized specification of QoS parameters as described in the present document. For specifications dealing with multiple services in separate clauses, the proposed structure is meant to be applied to each service clause. For specifications dealing with one service or topic only, the main clauses listed in the following may even be top-level clauses.

The specification should contain at least the following clauses:

- Introduction and overview
  Contains introductory text, phase diagrams, message sequence diagrams, etc.

- Points of observation
  Even for multi-service documents it might be preferable to have this clause on top-level instead of one for each service. For future extensibility, it should be present (as a placeholder) even if the specification currently does not need to add any PO to the ones defined by the present document. The structure of this clause should be like the one of clause 7 of the present document (see also clause B.3).

- Events
  Consists of one clause per abstract event (see clause B.4).

- QoS parameter definitions
  Consists of one clause "<phase name> phase" per identified phase plus additional clauses for QoS parameters not related to any phase, as far as applicable.

The structure of a phase clause should typically look like this:

- Phase definition
- <phase name> success
- <phase name> time
- … (additional QoS parameters related to this phase)

B.2 Top-down approach for deriving QoS parameter definitions

In the following, some guidance on the formal steps involved in defining QoS parameters for a particular service is given in the form of a short recipe:

- Divide the usage of the service into meaningful phases according to clause 6.1.
- For each phase:
  - Create a separate clause under "QoS parameter definitions" with a title like "<phase name> phase" (see clause B.1).
  - Identify the abstract events defining start, end and success of the phase.
  - Identify applicable contexts and corresponding POs, where the events identified in the previous step are supposed to be observed. If the definition of a PO does not exist yet, either create it locally, as a clause under "Points of observation" (see clauses B.1 and B.3), or have it added to the present document via CR.

    - If necessary, add appropriate event definition clauses under "Events" (see clauses B.1 and B.4).
    - Add the corresponding "phase definition" clause (see clause B.5.2).
    - If applicable, add the clauses for success and time (duration) of the phase (see clause B.6.1).
    - Identify any other QoS parameters related to this phase, together with applicable POs and possibly involved abstract events.
    - Create missing PO and/or event definitions.
    - Add a clause for each identified QoS parameter (see clause B.6).
- Identify QoS parameters not attributable to any phase and create separate clauses for them (see clause B.6) at the same level as the phase clauses, again after adding any missing PO and/or event definitions.

B.3 Specification of points of observation

B.3.1 Guidelines

Points of observation should be identified by the notation given at the end of clause 4.2. As a matter of convenience, this identifier should be used:

- as the title of the clause which specifies the PO
- as ID used in XML exports

PO definitions should contain a short description and, if applicable, refer to appropriate standard documents.

If possible, points of observation should be illustrated, e.g. by displaying their positions in protocol stack figures like the ones presented in clause 7.1. PO definitions may then refer to those illustrations informatively.

Clause 7 may serve as a reference on how to specify POs in general.

B.3.2 Specification template

The following template is proposed for a clause which specifies a particular point of observation (PO).

Usage: After copying the template, replace placeholders in angular brackets by the specific content they indicate and follow the instructions/hints given by blue meta-text. Finally, delete any blue meta-text parts.

Examples can be found in clause 7.

<clause number> <protocol or service access point acronym>@<reference point acronym> *(use appropriate heading style)*

ID	<protocol or service access point acronym>@<reference point acronym>
Description	<short description of the PO>
Related overview figure	<reference to overview figure> *or* n.a.
Related standards	<name of standard> [<number of normative reference>], etc.

B.4 Specification of events

B.4.1 Guidelines

There should be a separate clause for each identified abstract event. A short form of the abstract characterization should be used as clause title, e.g. "Telephony complete address information sent".

The clause should contain:

- an event ID suitable for XML exports;
- a description (long form of abstract characterization) of the event;
- a description of the user perception;
- an (extensible) list of specifications of technical event realizations for different applicable contexts and POs (according to clause 4.3.4).

A compact, tabular format for an event definition clause is proposed in clause B.4.2.

B.4.2 Specification template

The following template is proposed for a clause which specifies a particular event according to clause 4.3.

Usage: After copying the template, replace placeholders in angular brackets by the specific content they indicate and follow the instructions/hints given by blue meta-text. Finally, delete any blue meta-text parts.

<clause number> <event name> *(use appropriate heading style)*

Event ID	`<event ID>`	
Description	<extended abstract event description>	
User perception	<user action or observation>	
Context (e.g. access technology)	**PO**	**Technical description/Protocol event**
<context>	<PO name/ID> (see <reference to clause of PO definition>)	*e.g.* <protocol> <message> sent/received
<context>	<PO name/ID> (see <reference to clause of PO definition>)	*e.g.* <protocol> <message> sent/received (see note)
NOTE: <some note if applicable>		

B.5 Specification of phases

B.5.1 Guidelines

According to the guidelines outlined in clauses B.1 and B.2, QoS parameter definitions related to a particular phase of service usage should be grouped together under a common clause for that phase. In this common clause, the clauses defining QoS parameters should be preceded by a clause specifying the phase itself. This clause should always have the title "Phase definition".

A phase should be defined as a transaction in the sense of clause 4.4, i.e. by specifying which abstract events play the roles of its start, end and success triggers E_{start}, E_{end} and E_{success}, respectively.

Following the explanations in clause 4.4, the events themselves (and their realizations on different POs) should not be defined here, but under a separate "Events" clause for better reusability and extensibility (see clauses B.1, B.2 and B.4).

The specification of E_{start}, E_{end} and E_{success} should indicate the name (short form characterization) of the corresponding event and refer to the clause where the event is defined, e.g.:

EXAMPLE: E_{start} = "Telephony complete address information sent" (see clause …)
$E_{\mathrm{end}} = E_{\mathrm{success}}$ = "Telephony call set-up notification received" (see clause …)

In addition, the phase definition should contain an ID for the phase, suitable for XML exports. A compact table format for these items is proposed in clause B.5.2.

If necessary, preconditions for the phase and other explanations may be added as well.

B.5.2 Specification template

The following template is proposed for the clause "Phase definition" within a clause which specifies QoS parameters for a particular phase of service usage.

Usage: After copying the template, replace placeholders in angular brackets by the specific content they indicate and follow the instructions/hints given by blue meta-text. Finally, delete any blue meta-text parts.

<clause number> Phase definition *(use appropriate heading style)*

The <phase name> phase is defined by table <reference to the following phase definition table>.

In the following table, if $E_{end}=E_{success}$ delete the last two lines; otherwise delete the third line:

Table <table number>: <Phase name> - Phase definition

Phase ID	`<phase ID>`
Start trigger E_{start}	"<event name>" (see <reference to clause of event definition>)
End/success trigger $E_{\mathrm{end}} = E_{\mathrm{success}}$	"<event name>" (see <reference to clause of event definition>)
End trigger E_{end}	"<event name>" (see <reference to clause of event definition>)
Success condition E_{success}	E_{end} AND "<event name>" (see <reference to clause of event definition>)

B.6 Specification of QoS parameters

B.6.1 Guidelines

In addition to the mandatory requirements specified in clause 4.1, the specification of a QoS parameter should always include a short textual description of the measurable quantity for better readability.

EXAMPLE: *"The telephony call setup duration is the time between sending of complete address information and receipt of call set-up notification."*.

In order to be able to use QoS parameters in formulas in accordance with the ETSI drafting rules, appropriate symbols need to be introduced for them. Depending on the context it may be sufficient, in some cases, to use the generic symbol Q introduced in clause 4.1, indicating to which QoS parameter it refers in that particular context.

A more granular approach is to introduce different symbols for different types of QoS parameters (e.g. of different physical dimension). This has been done in clause 5 for some frequently used types. Further differentiation may be achieved by extending symbols with textual subscripts. Whenever new types of quantities are identified which may be used frequently, it should be considered to add them to clause 5 of the present document via CR.

It is recommended to make the distinction between quantity and aggregation required by clause 4.1 clearly visible by using fixed prefixes like "Measurable quantity:" and "Aggregated QoS parameter:".

The recommended way of presenting the formulas involved in a QoS parameter definition depends on how frequently formulas of the same type/structure occur:

- Specific formulas not used elsewhere should be fully specified according to the ETSI drafting rules (appropriately typeset formula followed by explicit explanation of any symbols occurring in that formula), like the formulas introduced in clause 5.
- For quantities of a type for which a generic formula exists elsewhere (e.g. in clause 5 of the present document), the formula itself may still be copied for better readability, but the full explanation should be replaced by a reference to the original definition, giving just a short indication about necessary substitutions. The same holds for the sample averages specified in clause 5. Examples for this style can be found in clauses 6.2 and 6.3.
- QoS parameters measuring success and time of a phase are so ubiquitous that they may be handled in a very compact way: It is suggested just to refer to clauses 6.2 and 6.3, respectively, of the present document, and to the respective "Phase definition" clause for the trigger events.

Apart from the (abstract) definition, the specification of a QoS parameter should indicate any special assumptions, preconditions and restrictions for applying the definition to particular contexts and POs.

B.6.2 Specification template

The present document currently does not suggest any particular template for the specification of a QoS parameter. Suitable templates are for further study and might be added in a future revision of the present document.

History

Document history		
V1.1.1	December 2018	Publication