International Telecommunication Union

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

G.1028

(04/2016)

SERIES G: TRANSMISSION SYSTEMS AND MEDIA, DIGITAL SYSTEMS AND NETWORKS

Multimedia Quality of Service and performance – Generic and user-related aspects

---

End-to-end quality of service for voice over 4G mobile networks

Recommendation ITU-T G.1028

ITU-T

ITU-T G-SERIES RECOMMENDATIONS

TRANSMISSION SYSTEMS AND MEDIA, DIGITAL SYSTEMS AND NETWORKS

INTERNATIONAL TELEPHONE CONNECTIONS AND CIRCUITS	G.100–G.199
GENERAL CHARACTERISTICS COMMON TO ALL ANALOGUE CARRIER-TRANSMISSION SYSTEMS	G.200–G.299
INDIVIDUAL CHARACTERISTICS OF INTERNATIONAL CARRIER TELEPHONE SYSTEMS ON METALLIC LINES	G.300–G.399
GENERAL CHARACTERISTICS OF INTERNATIONAL CARRIER TELEPHONE SYSTEMS ON RADIO-RELAY OR SATELLITE LINKS AND INTERCONNECTION WITH METALLIC LINES	G.400–G.449
COORDINATION OF RADIOTELEPHONY AND LINE TELEPHONY	G.450–G.499
TRANSMISSION MEDIA AND OPTICAL SYSTEMS CHARACTERISTICS	G.600–G.699
DIGITAL TERMINAL EQUIPMENTS	G.700–G.799
DIGITAL NETWORKS	G.800–G.899
DIGITAL SECTIONS AND DIGITAL LINE SYSTEM	G.900–G.999
MULTIMEDIA QUALITY OF SERVICE AND PERFORMANCE – GENERIC AND USER-RELATED ASPECTS	**G.1000–G.1999**
TRANSMISSION MEDIA CHARACTERISTICS	G.6000–G.6999
DATA OVER TRANSPORT – GENERIC ASPECTS	G.7000–G.7999
PACKET OVER TRANSPORT ASPECTS	G.8000–G.8999
ACCESS NETWORKS	G.9000–G.9999

For further details, please refer to the list of ITU-T Recommendations.

Recommendation ITU-T G.1028

End-to-end quality of service for voice over 4G mobile networks

Summary

Recommendation ITU-T G.1028 provides guidelines concerning the key aspects impacting end-to-end performance of managed voice applications over LTE networks and how they can be properly assessed using current elements of knowledge.

Some typical end-to-end scenarios are described, involving cases with LTE access at both sides of the communication, or with a different access technology at one side (wireless or wireline access). These scenarios are based on typical reference connections defined in this Recommendation, composed of various segments, including: terminal, wireless access, backhaul network, core network. Considerations regarding the sharing of the budget of some key parameters and the location where they can be assessed across these segments are provided.

History

Edition	Recommendation	Approval	Study Group	Unique ID*
1.0	ITU-T G.1028	2016-04-06	12	11.1002/1000/12748

* To access the Recommendation, type the URL http://handle.itu.int/ in the address field of your web browser, followed by the Recommendation's unique ID. For example, http://handle.itu.int/11.1002/1000/11830-en.

FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications, information and communication technologies (ICTs). The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of ITU. ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Assembly (WTSA), which meets every four years, establishes the topics for study by the ITU-T study groups which, in turn, produce Recommendations on these topics.

The approval of ITU-T Recommendations is covered by the procedure laid down in WTSA Resolution 1.

In some areas of information technology which fall within ITU-T's purview, the necessary standards are prepared on a collaborative basis with ISO and IEC.

NOTE

In this Recommendation, the expression "Administration" is used for conciseness to indicate both a telecommunication administration and a recognized operating agency.

Compliance with this Recommendation is voluntary. However, the Recommendation may contain certain mandatory provisions (to ensure, e.g., interoperability or applicability) and compliance with the Recommendation is achieved when all of these mandatory provisions are met. The words "shall" or some other obligatory language such as "must" and the negative equivalents are used to express requirements. The use of such words does not suggest that compliance with the Recommendation is required of any party.

INTELLECTUAL PROPERTY RIGHTS

ITU draws attention to the possibility that the practice or implementation of this Recommendation may involve the use of a claimed Intellectual Property Right. ITU takes no position concerning the evidence, validity or applicability of claimed Intellectual Property Rights, whether asserted by ITU members or others outside of the Recommendation development process.

As of the date of approval of this Recommendation, ITU had not received notice of intellectual property, protected by patents, which may be required to implement this Recommendation. However, implementers are cautioned that this may not represent the latest information and are therefore strongly urged to consult the TSB patent database at http://www.itu.int/ITU-T/ipr/.

Table of Contents

Recommendation ITU-T G.1028

End-to-end quality of service for voice over 4G mobile networks

1 Scope

This Recommendation works on the assumption that Voice over LTE (VoLTE) is a so-called "managed" voice service, in opposition to over the top (OTT) applications without use of session initiation protocol/IP multimedia subsystem (SIP/IMS) signalling and with no prioritized traffic. Video-telephony over LTE (ViLTE) is another service which will be addressed in a specific Recommendation.

This Recommendation describes the key aspects impacting end-to-end performance of managed voice applications over LTE networks, for the most common call cases (IMS tromboning is not considered, nor single radio voice call continuity (SRVCC), nor mobility under wireless local access network), and how they can be properly assessed using current elements of knowledge.

Relevant quality of service (QoS) mechanisms used to manage the voice service, such as robust header compression (RoHC), transmission time interval (TTI) bundling, semi-persistent scheduling (SPS), discontinuous transmission (DTX) and reception (DRX), service domain selection (SDS) or SIP preconditions, are not considered in this Recommendation as a mandatory part of a VoLTE service, but their impact on end-to-end perceived quality will be taken into account.

Analysis of the impact of VoLTE on quality of supplementary services (such as data streaming) or on device features (battery life) is outside the scope of this Recommendation.

2 References

The following ITU-T Recommendations and other references contain provisions which, through reference in this text, constitute provisions of this Recommendation. At the time of publication, the editions indicated were valid. All Recommendations and other references are subject to revision; users of this Recommendation are therefore encouraged to investigate the possibility of applying the most recent edition of the Recommendations and other references listed below. A list of the currently valid ITU-T Recommendations is regularly published. The reference to a document within this Recommendation does not give it, as a stand-alone document, the status of a Recommendation.

[ITU-T E.800] Recommendation ITU-T E.800 (2008), *Definitions of terms related to quality of services*.

[ITU-T E.804] Recommendation ITU-T E.804 (2014), *Quality of service aspects for popular services in mobile networks*.

[ITU-T G.107] Recommendation ITU-T G.107 (2015), *The E-model: a computational model for use in transmission planning*.

[ITU-T G.107.1] Recommendation ITU-T G.107.1 (2015), *Wideband E-model*

[ITU-T G.109] Recommendation ITU-T G.109 (1999), *Definition of categories of speech transmission quality*.

[ITU-T G.114] Recommendation ITU-T G.114 (2003), *One-way transmission time*

[ITU-T G.711] Recommendation ITU-T G.711 (1988), *Pulse code modulation (PCM) of voice frequencies*.

[ITU-T G.1000] Recommendation ITU-T G.1000 (2001), *Communications quality of service: A framework and definitions*.

[ITU-T P.10] Recommendation ITU-T P.10/G.100 (2006), *Vocabulary for performance and quality of service*.

[ITU-T P.563] Recommendation ITU-T P.563 (2004), *Single-ended method for objective speech quality assessment in narrow-band telephony applications*.

[ITU-T P.564] Recommendation ITU-T P.564 (2007), *Conformance testing for voice over IP transmission quality assessment models*.

[ITU-T P.800.1] Recommendation ITU-T P.800.1 (2016), *Mean opinion score (MOS) terminology*.

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

[ITU-T P.863] Recommendation ITU-T P.863 (2014), *Perceptual objective listening quality assessment*.

[ITU-T P.863.1] Recommendation ITU-T P.863.1 (2014), *Application guide for Recommendation ITU-T P.863*.

[ITU-T Y.1540] Recommendation ITU-T Y.1540 (2016), *Internet protocol data communication service – IP packet transfer and availability performance parameters*.

[ITU-T Y.1541] Recommendation ITU-T Y.1541 (2011), *Network performance objectives for IP-based services*.

[3GPP TS 22.105] 3GPP TS 22.105 v12.1.0 (2014), *3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Services and service capabilities (Release 12)*.

[3GPP TS 23.203] 3GPP TS 23.203 v 13.5.1 (2015), *3rd Generation Partnership Project*; *Technical Specification Group Services and System Aspects*; *Policy and charging control architecture (Release 13)*.

[3GPP TS 24.229] 3GPP TS 24.229 v 13.3.1 (2015), *3rd Generation Partnership Project*; *Technical Specification Group Core Network and Terminals*; *IP multimedia call control protocol based on Session Initiation Protocol (SIP) and Session Description Protocol (SDP); Stage 3 (Release 13)*.

[3GPP TS 26.114] 3GPP TS 26.114 v 13.1.0 (2015), *3rd Generation Partnership Project*; *Technical Specification Group Services and System Aspects*; *IP Multimedia Subsystem (IMS); Multimedia Telephony; Media handling and interaction (Release 13)*.

[3GPP TS 26.131] 3GPP TS 26.131 v 13.1.0 (2015), *3rd Generation Partnership Project*; *Technical Specification Group Services and System Aspects*; *Terminal acoustic characteristics for telephony; Requirements (Release 13)*.

[ETSI TS 101 563] ETSI TS 101 563 v 1.4.1 (2015), *Speech and multimedia Transmission Quality* (STQ); *IMS/PES/VoLTE exchange performance requirements*.

[ETSI TR 103 219] ETSI TR 103 219 v1.1.1 (2015), *Speech and multimedia Transmission Quality* (STQ); *Quality of Service aspects of voice communication in an LTE environment*.

[IETF RFC 6076] IETF RFC 6076 (2011), *Basic Telephony SIP End-to-End Performance Metrics*.

[IETF RFC 3095] IETF RFC 3095 (2001), *Robust Header Compression (RoHC): Framework and four profiles: RTP, UDP, ESP, and uncompressed*.

3 Definitions

3.1 Terms defined elsewhere

This Recommendation uses the following terms defined elsewhere:

3.1.1 call set-up time [ITU-T E.800]: The period starting when the address information required for setting up a call is received by the network (recognized on the calling user's access line) and finishing when the called party busy tone, or ringing tone or answer signal is received by the calling party (i.e., recognized on the calling user's access line). Local, national and service calls should be included, but calls to Other Licensed Operators should not, as a given operator cannot control the QoS delivered by another network.

3.1.2 speech quality [ITU-T P.10]: Quality of spoken language as perceived when acoustically displayed. Result of a perception and assessment process, in which the assessing subject establishes a relationship between the perceived characteristics, i.e., the auditory event, and the desired or expected characteristics.

3.1.3 service accessibility performance [ITU-T E.800]: The ability of a service to be obtained, within specified tolerances and other given conditions, when requested by the user.

3.1.4 mean one-way propagation time [ITU-T P.10]: In a connection, the mean of the propagation times in the two directions of transmission.

3.2 Terms defined in this Recommendation

None.

4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

2G	Second Generation of radio access network (also called GSM)
3G	Third Generation of radio access network (also called UMTS)
4G	Fourth Generation of radio access network (also called LTE)
ACK	Acknowledgment
AEC	Acoustical Echo Cancellation
AF	Application Function
AGC	Automatic Gain Control
AGW	Access Gateway
AMR	Adaptive Multi Rate coding
AMR-WB	Wide-Band Adaptive Multi Rate coding
BGCF	Border Gateway Control Function
BTS	Base Transceiver Station
CDR	Call Detail Record
CNG	Comfort Noise Generation
CS	Circuit Switched

CSFB	Circuit Switched Fall Back
DL	Downlink
DRX	Discontinuous Reception
DTX	Discontinuous Transmission
EEC	Electrical Echo Cancellation
E-NodeB	Enhanced Node B
EPC	Evolved Packet Core
E-UTRAN	Evolved UMTS Terrestrial Radio Access Network
EVS	Enhanced Voice Services
GBR	Guaranteed Bit Rate
GERAN	GSM/Edge Radio Access Network
GGSN	Gateway GPRS (General Packet Radio Service) Service Node
GSM	Global System for Mobile Communications
GSMA	The GSM Association
GW	Gateway
HARQ	Hybrid Automatic-Repeat-Request
HD voice	High Definition voice
HO	Hand Over
HSS	Home Subscriber Server
I-CSCF	Interrogating Call Session Control Function
IMS	IP Multimedia Subsystem
IPDV	Internet Packet Delay Variation
IPER	Internet Packet Error Ratio
IPLR	Internet Packet Loss Ratio
IPTD	Internet Packet Transfer Delay
IRA	Ineffective Registration Attempt
KPI	Key Performance Indicator
LTE	Long Term Evolution (of radio access networks)
MGCF	Media Gateway Controller Function
MGW	Media Gateway
MME	Mobility Management Entity
MOS	Mean Opinion Score
MOS-LQ	Mean Opinion Score – Listening Quality
MTAS	Multimedia Telephony Application Server
MTSI	Multimedia Telephony Service for IMS
NB	Narrow Band (voice spectrum)
NER	Network Error Rate

O-SDS	Originating Service Domain Selection
OTT	Over-The-Top
PCC	Policy and Charging Control
PCEF	Policy and Charging Enforcement Function
PCRF	Policy and Charging Rule Function
P-CSCF	Proxy Call Session Control Function
PDD	Post Dialling Delay
PGW	Packet Data Network Gateway
PLC	Packet Loss Concealment
PLMN	Public Land Mobile Network
PSTN	Public Switched Telephone Network
QCI	Quality Classification Identifier
QoS	Quality of Service
RACH	Random Access Channel
RoHC	Robust Header Compression
RRC	Radio Resource Control
RTCP	Real-time Transport Control Protocol
RTP	Real-time Transport Protocol
S-CSCF	Serving Call Session Control Function
SCR	Session Completion Ratio
SDP	Session Description Protocol
SDS	Service Domain Selection
SEER	Session Establishment Effectiveness Ratio
SGSN	Service GPRS (General Packet Radio Service) Support Node
SGW	Serving Gateway
SIP	Session Initiation Protocol
SIP-I	Session Initiation Protocol with encapsulated ISUP
SPR	Subscriber Profile Repository
SPS	Semi-Persistent Scheduling
SRD	Session Request Delay
SRVCC	Single Radio Voice Call Continuity
SWB	Super Wide Band (voice spectrum)
T-ADS	Terminating Access Domain Selection
TDM	Time Domain Multiplication
TrFO	Transcoding Free Operation
TrGW	Trunking Gateway
T-SDS	Terminating Service Domain Selection

TTI	Transmission Time Interval
UDP	User Datagram Protocol
UE	User's Equipment
UL	Uplink
UMTS	Universal Mobile Telecommunication System
UTRAN	UMTS Terrestrial Radio Access Network
VAD	Voice Activity Detection
ViLTE	Video-telephony over LTE
VoWiFi	Voice over WiFi
VoLTE	Voice over LTE
VQE	Voice Quality Enhancement
WB	Wide Band (voice spectrum)
WiFi	Wireless Fidelity radio access network

5 Conventions

None.

6 Brief introduction on Voice over LTE and assumptions

This Recommendation makes the following assumptions, in line with the IMS profile for voice defined by the global system for mobile communications association (GSMA) in [b-GSMA IR.92] and the multimedia telephony service for IMS (MTSI) media handling procedures (voice part only) defined by 3GPP in [3GPP TS 26.114], as far as Voice over LTE (VoLTE) is concerned:

– Voice calls are initiated and/or received by a terminal connected to a 4G radio access network (evolved universal terrestrial radio access network, or E-UTRAN).
– Terminals connected to E-UTRAN embed a VoLTE client by default, either directly native on the chip or as a dedicated application.
– E-UTRAN is connected to an evolved packet core (EPC) network. The connecting point is an enhanced node B (e-NodeB).
– EPC enables the communication between 4G terminals and IMS core platforms, in order to set up calls using the session initiation protocol (SIP). There is no other possibility to set up calls.
– EPC supports quality of service (QoS) classification (between policy and charging enforcement function (PCEF) and the terminal), as defined in [3GPP TS 22.105] (section 5) and [3GPP TS 23.20]3 (section 6.1.7 and associated table partially reproduced in Table 1).
– SIP signalling is assigned with the QoS quality classification identifier (QCI) 5.
– Real time voice signal during calls is carried out with the real-time transport protocol/user datagram protocol (RTP/UDP), and is assigned with the QCI 1.
– The application of QCI 5 for SIP signalling bearer and QCI 1 for real time voice bearer gives them a high priority over all data bearer, resulting in no degradation of the voice service when used in parallel with data session on the same device.
– Voice calls are possible seamless with any far end user of circuit switched (CS) voice connected to public switched telephone network (PSTN) or to 2G or 3G mobile access. This

implies that the VoLTE service is assumed being deployed over the entire 4G network. Cases of VoLTE terminals being under 4G but non-VoLTE coverage are not considered.

– Voice calls between two VoLTE clients can be placed using any bit rate and codec mode of adaptive multi rate coding (AMR) and wide-band adaptive multi rate coding (AMR-WB), as indicated in section 3.2.1 of [b-GSMA IR.92]. In the near future, the super wide band (SWB) codec enhanced voice services (EVS) will also be supported. Entities ending the user plane (terminal, media gateways (MGWs)) support transcoding free operation (TrFO).

– The feature TrFO on SIP session initiation protocol with encapsulated ISUP (SIP-I) is activated which means that codec negotiation between VoLTE and the other networks is performed in an end-to-end way.

Table 1 – Quality classification identifiers in use for voice over LTE (source: [3GPP TS 23.203])

QCI	Resource type	Priority level	Packet delay budget	Packet error loss rate	Example services
1	GBR	2	100 ms (Note)	10^{-2}	Conversational voice
5	Non-GBR	1	100 ms (Note)	10^{-6}	IMS signalling
NOTE – A delay of 20 ms for the delay between a PCEF and a radio base station should be subtracted from a given packet delay budget to derive the packet delay budget that applies to the radio interface. This delay is the average between the case where the PCEF is located "close" to the radio base station (roughly 10 ms) and the case where the PCEF is located "far" from the radio base station, e.g., in case of roaming with home routed traffic (the one-way packet delay between Europe and North America west coast is roughly 50 ms). The average takes into account that roaming is a less typical scenario. It is expected that subtracting this average delay of 20 ms from a given packet delay budget will lead to desired end-to-end performance in most typical cases. Also, note that the packet delay budget defines an upper bound. Actual packet delays – in particular for guaranteed bit rate (GBR) traffic – should typically be lower than the packet delay budget specified for a QCI as long as the terminal has sufficient radio channel quality.					

These elements are constitutive of a so-called "managed" voice service, in opposition to OTT applications without use of SIP/IMS signalling and with no prioritized traffic.

Relevant QoS mechanisms used in 4G access network or in EPC and IMS platforms to manage the voice service are not considered in this Recommendation as a mandatory part of a VoLTE service, but their impact on end-to-end perceived quality will be taken into account as far as possible. The most frequent of these features are:

– Robust header compression (RoHC, see [IETF RFC 3095]) is a method to shorten the size of voice packets (header often represents 60% of the total of a packet, and can be reduced by a factor of up to 20 thanks to RoHC) and thus limit the network bandwidth occupation.

– Transmission time interval (TTI) bundling is a mechanism to improve coverage at 4G cell edge or in poor radio conditions, by transmitting each packet four times in a row (without signalling overhead).

– Semi-persistent scheduling simplifies the allocation and re-allocation of radio resources in a cell and thus can allow more simultaneous VoLTE calls and a lower consumption of battery in devices.

– Discontinuous transmission (DTX) is a technique to reduce the bandwidth occupied by a voice call, by not transmitting data during silent periods.

– Discontinuous reception (DRX) puts the terminal into periodic repetition of sleep mode and wake up mode (to listen to messages from the network), allowing saving battery consumption.

– SIP preconditions, as specified in section 2.4.1 of [b-GSMA IR.92] and [3GPP TS 24.229], can be added to session description protocol (SDP) messages exchanged between terminals and IMS in order to ensure that the dedicated radio bearers are available before setting up the call. This strengthens the call set-up process but makes it longer by up to several seconds.

– Terminating or originating service domain selection (T-SDS and O-SDS) allows selecting between the CS or the IMS domain to provide the voice service to a VoLTE terminal under VoLTE and CS network coverage. The mechanism for the selection of domain to communicate call termination from the network to the terminal is called terminating access domain selection (T-ADS).

7 Hypothetical reference models

The global architecture supporting VoLTE services can be seen in Figure 1:

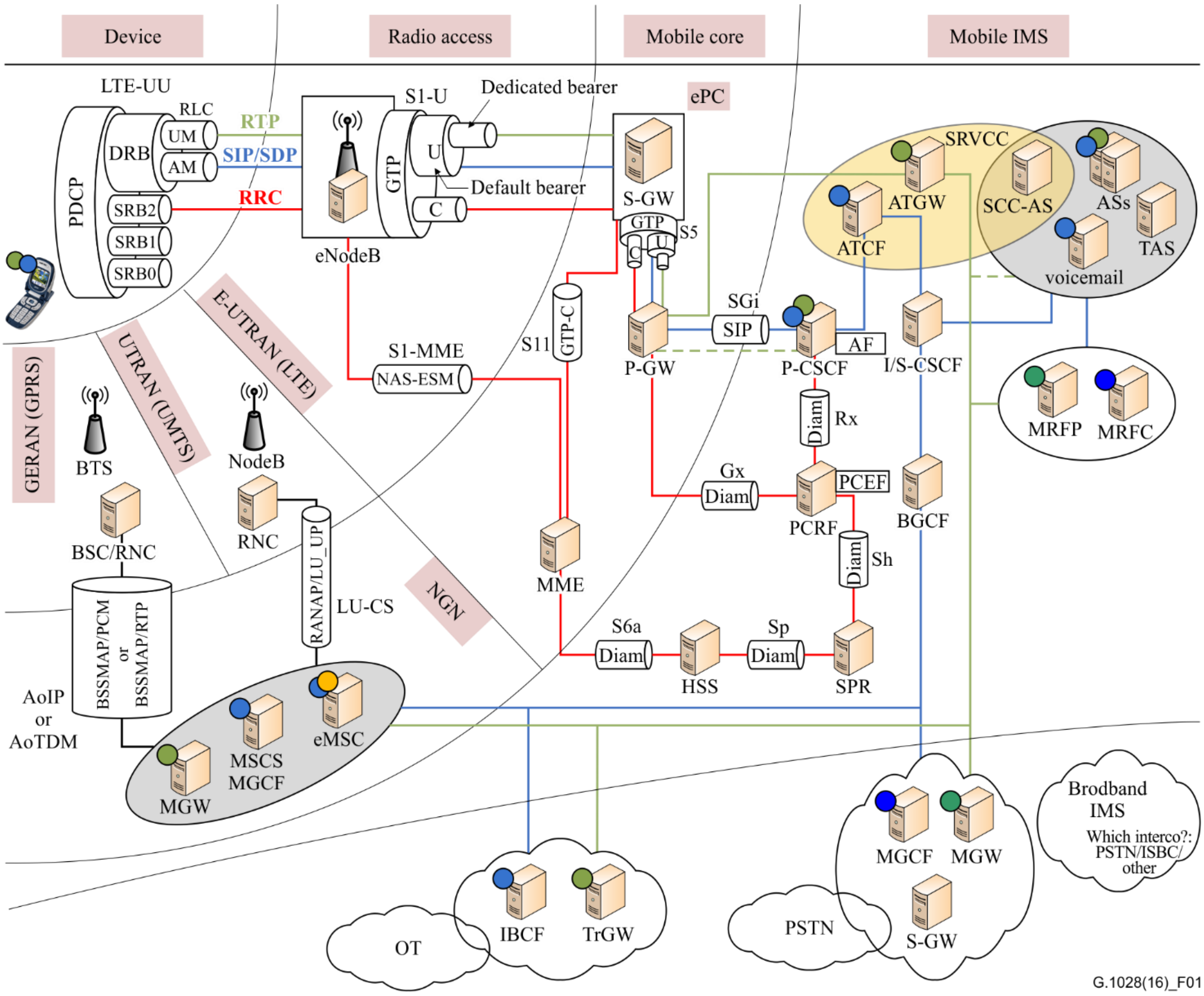

Figure 1 – Overall architecture of VoLTE services

Several blocks are present and form the basic elements of a reference model:

– The terminal

– The E-UTRAN

– The EPC network
– The mobile IMS core network
– Other radio access networks (GERAN and UTRAN)
– Mobile core networks for 2G and 3G
– Interconnection with other EPC or fixed networks

The description above is simplified in Figure 2:

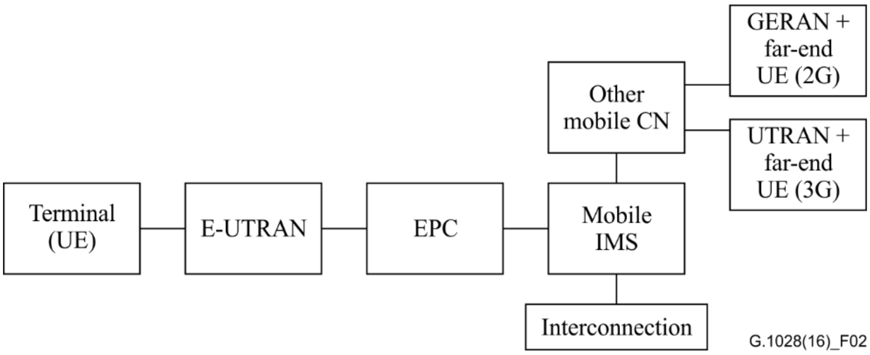

Figure 2 – Simplified architecture of VoLTE services

This Recommendation provides an overview of the impact of various issues on perceived quality, together with an estimation of the quantization of this impact by building block of the hypothetical reference model, for several end-to-end scenarios falling into the scope of this Recommendation.

A client of a VoLTE service can experience different types of calls:

– Basic call: Either with another VoLTE user connected to the same 4G network (see clause 7.1) or with a user of another voice network (CS or PSTN, see clauses 7.3 and 7.4).
– Circuit switched fall back (CSFB): A call with another 4G terminal, when one of the two ends must perform a fall back to a CS connection over 3G or 2G before call set up. From a user point of view, CSFB is automatic and transparent, no action is required. The performance of CSFB in terms of call set up is seen as a sub-part of basic call performance. Once the CSFB is performed, the performance objectives in terms of integrity and call retainability are similar to the ones of a basic call (see clauses 7.3 and 7.4).
– Interconnection: A call between two VoLTE terminals connected to two different interconnected networks (see clause 7.2).
– IMS tromboning: When the VoLTE terminal is under CS coverage, signalling and user planes go through the IMS domain. From the e2e performance perspective, this IMS tromboning should only impact end-to-end delay and post dialling delay (PDD). This call case is outside the scope of the present Recommendation.

Due to mobility, a call initiated under a 4G VoLTE coverage may have to hand over (HO) to CS coverage in order to continue. This process, known as SRVCC, is also outside of the current scope of this Recommendation and under consideration for a further revision.

A call initiated under a 4G VoLTE coverage may also have to hand over to wireless local access network coverage. A 4G terminal may also directly start a voice call on an IMS platform under this radio coverage. This use case, known as voice over WiFi (VoWiFi), is also outside of the current scope of this Recommendation and under consideration for a further revision.

The most common scenarios, considered in this Recommendation, are detailed below.

7.1 LTE-LTE communication on the same network

As shown in Figure 3, two clients of the same VoLTE service are in communication. It is assumed that they are not located under the coverage of the same cell.

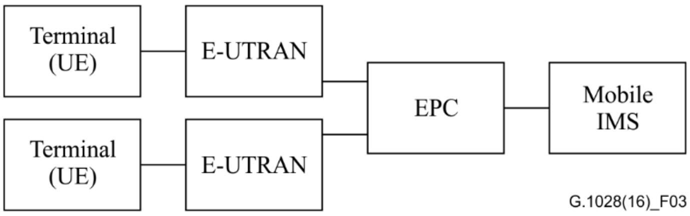

Figure 3 – Hypothetical reference model for an LTE-LTE communication on the same network

7.2 LTE-LTE communication between two interconnected networks (includes roaming)

Figure 4 shows clients of VoLTE services of two different operators in communication. The IMS core networks are interconnected, allowing both signalling and user plane continuity.

It is assumed that this interconnection is IP-based (time domain multiplication (TDM) solution is not considered). Also TrFO on SIP SIP-I feature is assumed activated on both networks.

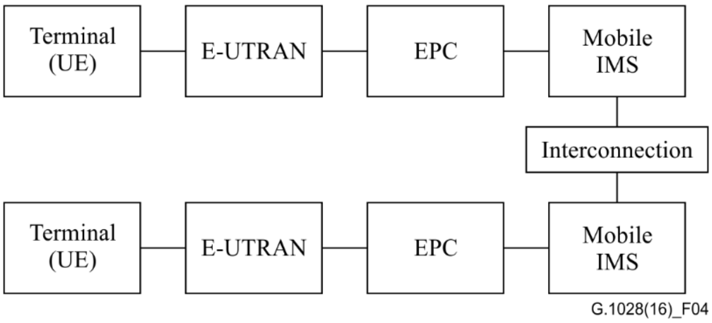

Figure 4 – Hypothetical reference model for an LTE-LTE communication between two interconnected networks

Two different VoLTE interconnections cases can be addressed:

– The two networks are directly interconnected which yields the use of the same codec – as for intra public land mobile network (PLMN) case – and low values of PDD and end-to-end delay.

– The two networks are interconnected through a CS intermediate network that increases PDD and end-to-end delay, reduces the codec rate, and might even cause the loss of high definition voice (HD voice) (if WB AMR codec is not supported by the CS network).

7.3 LTE-3G communication

When a VoLTE client is in communication with a CS mobile client, the mobile IMS is connected to the 3G mobile core network (see Figure 5), allowing both signalling and user plane continuity.

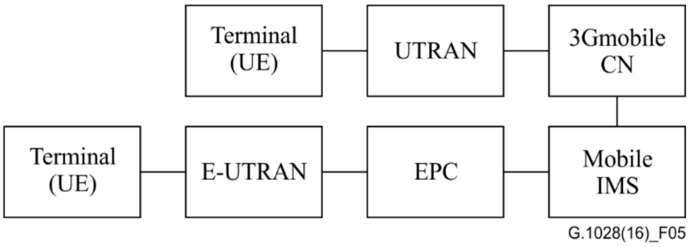

Figure 5 – Hypothetical reference model for an LTE-3G communication

7.4 LTE-PSTN communication

When a VoLTE client is in communication with a user on PSTN, the mobile IMS is connected to the fixed IMS (see Figure 6), allowing both signalling and user plane continuity.

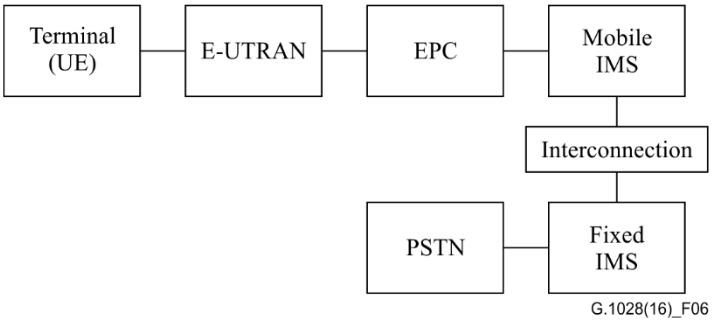

Figure 6 – Hypothetical reference model for an LTE-PSTN communication

All scenarios including interconnection to another network may include very long international paths. This is considered as a separate case and does not lead to considerations for the general budgeting of delays.

8 Aspects of quality of service for voice over LTE services

Figure 7 shows the QoS degradations that can be typically encountered on a VoLTE call. QoS is understood here as defined in [ITU-T E.800]. The main elements of the network are depicted to show the signalling and media elements as well as the connections with PSTN or mobile platforms.

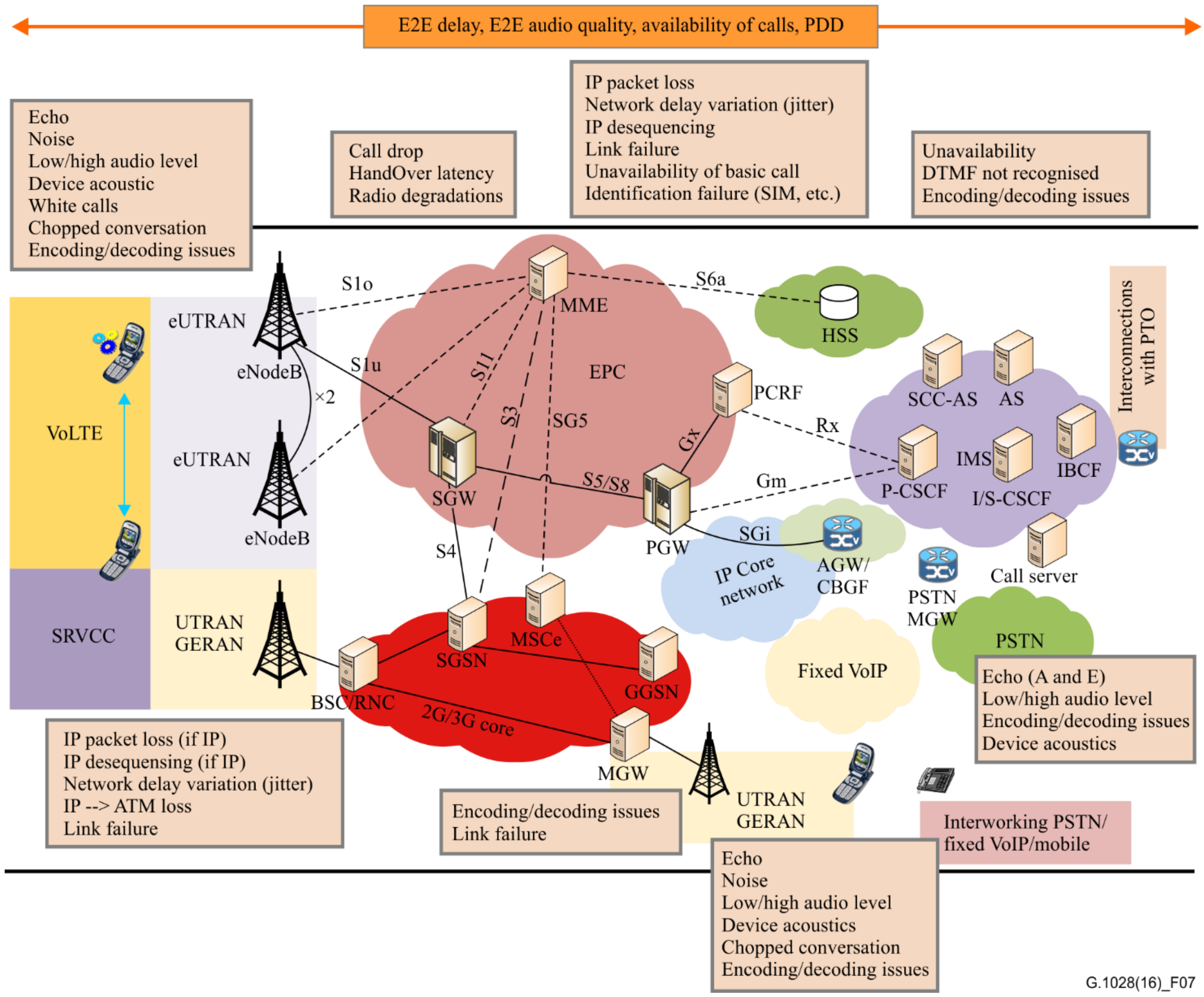

Figure 7 – Typical degradations of VoLTE communications

From a customer point of view (QoS required and perceived, as defined in [ITU-T G.1000]), these degradations are divided into the following categories:

- Call session performance
  - Problems of registration to the service (IMS/SIP).
  - Call set up issues (bad accessibility).
  - Failed continuity (or retainability), including impact of mobility (radio hand-overs and SRVCC events).
- Perceived speech quality during the call (integrity)
  - Frequency content. This refers to the speech spectrum of signals presented to end users (NB, WB or SWB) and its potential distortions.
  - Interruptions. Concerns all events resulting in clipping of the speech signal during the conversation.
  - End to end delay (impact on conversation interactivity)
  - Presence of unwanted noises, from whatever origin.

A more detailed list of typical degradations encountered by end-users of VoLTE service, together with potential causes, is provided in Annex A.

VoLTE can also have an impact on other services or functionalities of the device. For instance:

– data transfer performance is potentially influenced by the presence of VoLTE usage on the same cell;

– battery autonomy can become critical if the usage of a given application (here VoLTE) is consuming too much power.

These aspects are however outside the scope of the present Recommendation.

To report, monitor and troubleshoot these QoS issues experienced by customers, consistent measurements architecture must be set up with appropriate metrics. This is the purpose of following sections.

9 Relevant indicators and quality targets

From an end-to-end perspective, as perceived by an end-user located under a VoLTE compliant 4G network coverage and originating calls, the relevant indicators are given in the Table 2, together with the potential contributor network key performance indicator(s) (KPI(s)).

Table 2 – End-to-end quality indicators and corresponding network KPIs

End-to-end indicators	Definition	IP network KPIs
Registration success rate	Rate of successful registration attempts in the VoLTE service Equivalent to IMS registration success ratio as defined in [ETSI TR 103 219]	<u>Registration success rate</u> KPI related to IMS and based on P-CSCF counters. Equivalent to 1 – ineffective registration attempt (IRA) ratio, as defined in [IETF RFC 6076]
Service availability	End to end service availability in terms of capacity to establish calls from, and to, a VoLTE customer. Equivalent to 1 – VoLTE session set-up failure ratio as defined in [ETSI TR 103 219] Equivalent to 1 – telephony service non-accessibility as defined in [ITU-T E.804] (clause 7.3.6.1)	<u>Network efficiency ratio</u> Measures the ability of network, from the service platform point of view, to deliver calls to the VoLTE customer Based on SIP protocol, network error rate (NER) is equivalent to session establishment effectiveness ratio (SEER), as defined in [IETF RFC 6076]
Post dialling delay	Time interval (in seconds) between the end of dialling by the caller and the reception back by him of the appropriate ringing tone or recorded announcement. Equivalent to call set-up time, as defined in [ITU-T E.800] Equivalent to telephony set-up time as defined in [ITU-T E.804] (clause 7.3.6.2)	<u>SIP session set-up time</u> Interval between sending INVITE message (with SDP) and ACK (180 or 200) message by the originating side. Equivalent to successful session request delay (SRD), as defined in [IETF RFC 6076]

Table 2 – End-to-end quality indicators and corresponding network KPIs

End-to-end indicators	Definition	IP network KPIs
Voice quality (MOS-LQ)	Equivalent to speech quality as defined in [ITU-T P.10]. Models like those defined in [ITU-T P.862] and [ITU-T P.863] provide an objective view on the quality of the voice signal as it may be perceived by the customer. Can be seen on a call basis or on a sample basis (see [ITU-T E.804] clauses 7.3.6.3 and 7.3.6.4)	Network quality index ([ITU-T G.107], [ITU-T P.564]) IP Packet loss ratio (see definition of Internet packet loss ratio (IPLR) in [ITU-T Y.1540]): several possible measurement points.
Mouth-to-ear delay	The time it takes for the speech signal to go from the mouth of the speaker to the ear of the listener	IP packet transfer delay (see definition of Internet packet transfer delay (IPTD) in [ITU-T Y.1540]) Round trip time Corresponds approximately to twice the end-to-end delay Can be measured based on RTCP protocol messages
Call drop rate	Service continuity in terms of capacity to maintain calls to their normal end. Equivalent to telephony cut-off call ratio as defined in [ITU-T E.804] (clause 7.3.6.5)	Session completion rate KPI related to IMS and based on P-CSCF counters Equivalent to session completion ratio (SCR), as defined in [IETF RFC 6076]
Speech bandwidth (NB, WB or SWB)	Measurement of the bandwidth used (normal NB or WB, or even partial and unwanted bandwidth limitation).	Codec statistics Information related to the selection of (AMR and AMR WB) codec and codec modes, as well as switch between them, accessible on SIP protocol messages.

Measurement points in VoLTE session set-up and completion procedures have been captured in Figures 8 and 9:

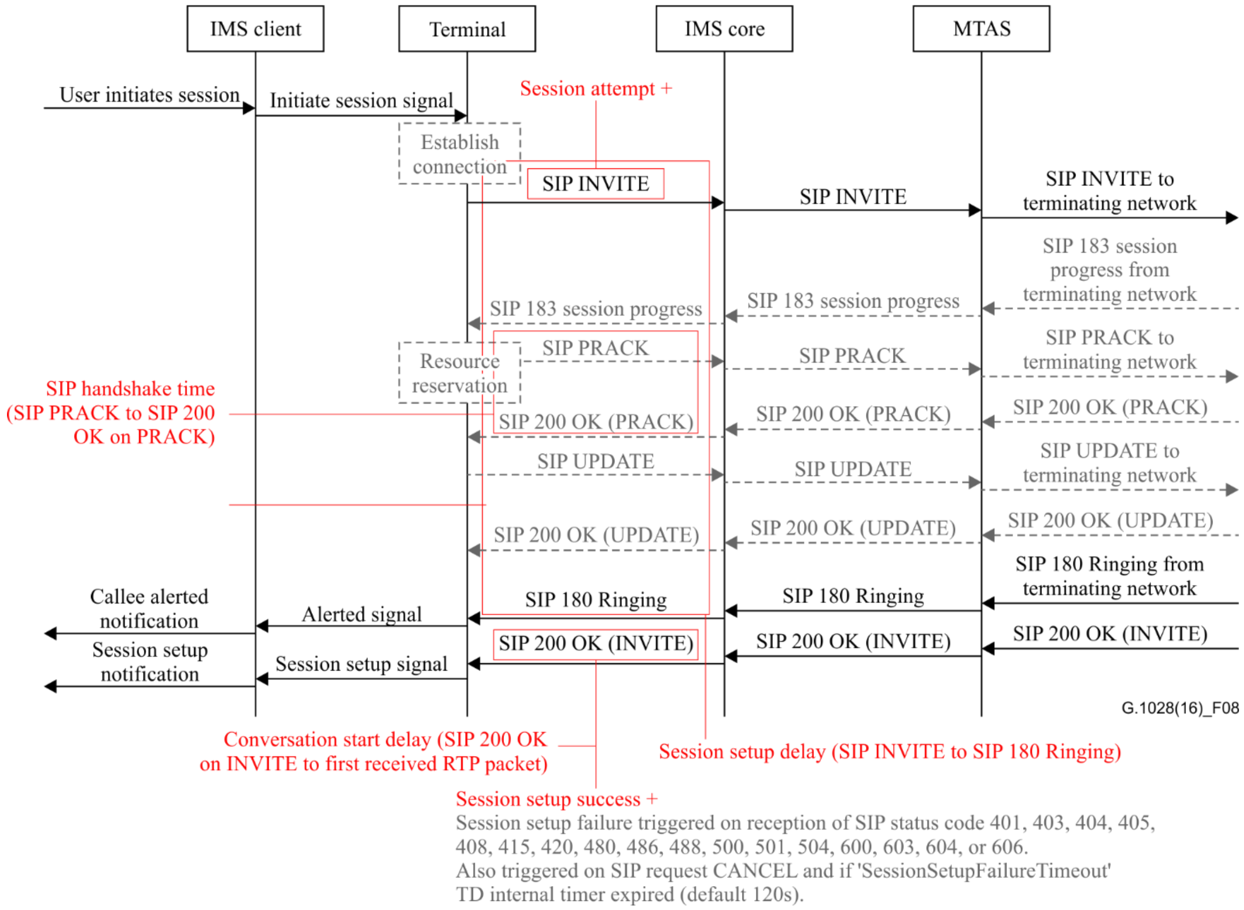

Figure 8 – VoLTE session set-up measurement flow

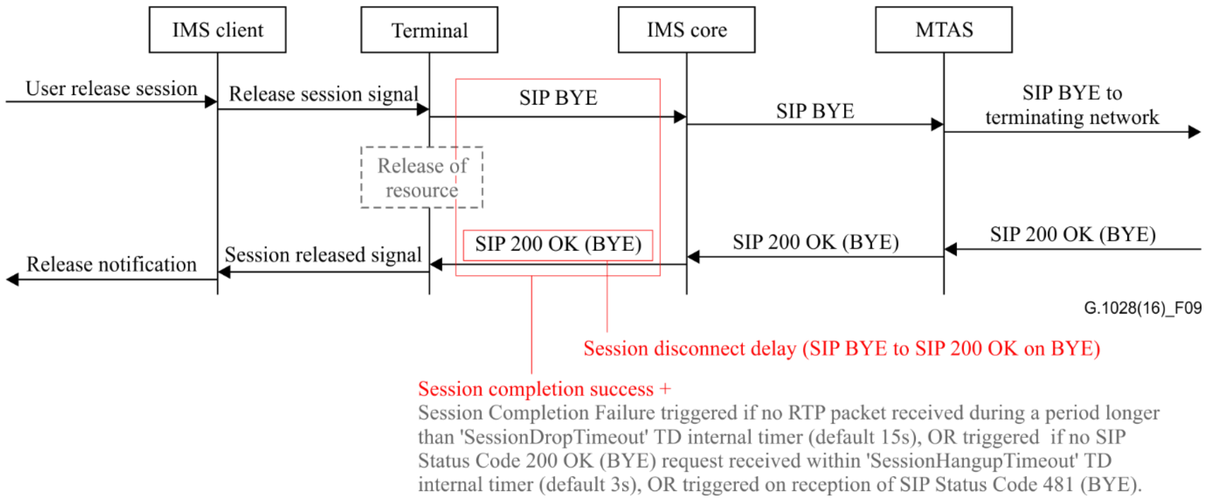

Figure 9 – VoLTE session completion measurement flow

For most of the indicators in Table 2, a budget can be assigned to the various segments that compose end-to-end paths as seen in clause 6. Tables 3 to 6 provide indications of target values that can be reasonably reached on each of these segments for each of the hypothetical reference connections depicted in clause 7. The total budget is not necessarily the exact sum of all individual budgets.

These targets are examples of realistic values that network operators may reach when using tools complying with up-to-date standards. For instance, the mean opinion scores (MOS) in Tables 3 to 6 are meant as average values when applying [ITU-T P.863] with the right reference sentence (i.e., complying with [ITU-T P.863.1]) and doing a small drive test with state-of-the art devices. For longer drive test, with e.g., hand over (HO) events, then these values may be seen as a maximum. For lab tests with clean stable environment, then this may be seen as minimum value. It should also be noted that this could change in the future with new devices and network set-ups.

LTE-LTE communication on the same network

Table 3 – Quality budgets for LTE-LTE communication on the same network

End-to-end indicators	**TOTAL budget**	**Terminal**	**E-UTRAN**	**EPC**	**Mobile IMS**	**Transmission network**
Registration success rate	99.9%	99.9%	99.9%	100%	99.9%	No target
Service availability	99% Note 1	No target	No target	No target	No target	No target
Post dialling delay (PDD)	LTE-LTE: 3.5 s (Note 2) CSFB: 6 s (Note 3)	No target	No target	No target	No target	No target
Voice quality (MOS-LQxSW)	4 (Note 4)	No target	No target	No target	No target	No target
Mouth-to-ear delay	400 ms (Note 5)	190 ms (sending + receiving) (Note 6)	80 ms on both sides	50 ms	0	10 ms (may be bigger for large countries)
Call drop rate	2 %	No target	No target	No target	No target	No target

NOTE 1 – Call processing performance objective according to [ETSI TS 101 563] is higher than 99.9%.
NOTE 2 – [ETSI TS 101 563] recommends 5.9 s, with 95% of probability below 2.4 s.
NOTE 3 – Only circuit switched fall back on mobile originating side is considered here.
NOTE 4 – Assumes the use of AMR-WB at a 23.85 kbit rate. The use of other codecs and/or bit rates will result in other values.
NOTE 5 – [ITU-T G.114] specifies a preferred maximum value at 150 ms, impossible to reach currently; some network operators are able to provide national calls with delays below 250 ms.
NOTE 6 – According to [3GPP TS 26.131].

LTE-LTE communication between two interconnected networks (includes roaming)

Table 4 – Quality budgets for LTE-LTE communication between two interconnected networks

End-to-end indicators	**TOTAL budget**	**Terminal**	**E-UTRAN**	**EPC**	**Mobile IMS**	**(International) transmission network**
Registration success rate	99.9%	99.9%	99.9%	100%	99.9%	No target
Service availability	99%	No target	No target	No target	No target	No target

Table 4 – Quality budgets for LTE-LTE communication between two interconnected networks

End-to-end indicators	TOTAL budget	Terminal	E-UTRAN	EPC	Mobile IMS	(International) transmission network
Post dialling delay (PDD)	LTE-LTE: 4 s CSFB: 6 s (Note 1)	No target	No target	No target	No target	No target
Voice quality (MOS-LQxSW)	4 (if HD voice + TrFO) (Note 2) 2.8 (otherwise)	No target	No target	No target	No target	No target
Mouth-to-ear delay	400 ms (Note 3)	190 ms (sending + receiving) (Note 4)	80 ms on both sides	50 ms	0	see typical values in [b-GSMA IR.34] Table 7
Call drop rate	2 %	No target	No target	No target	No target	No target

NOTE 1 – Only circuit switched fall back on mobile originating side is considered here.
NOTE 2 – Assumes the use of AMR-WB at a 23.85 kbit rate. The use of other codecs and/or bit rates will result in other values.
NOTE 3 – [ITU-T G.114] specifies a preferred maximum value at 150 ms, impossible to reach currently; some network operators are able to provide national calls with delays below 250 ms.
NOTE 4 – According to [3GPP TS 26.131].

LTE-3G communication

Table 5 – Quality budgets for LTE-3G communication

End-to-end indicators	TOTAL budget	Terminal	E-UTRAN	EPC	Mobile IMS	(Inter-national) trans-mission network	Far end access network	Far end terminal
Registration success rate	99.9%	99.9%	99.9%	100%	99.9%	No target	No target	No target
Service availability	98%	No target	No target	No target	No target	No target	No target	No target
Post dialling delay (PDD)	4.5 s CSFB: for further study	No target	No target	No target	No target	No target	No target	No target

Table 5 – Quality budgets for LTE-3G communication

End-to-end indicators	TOTAL budget	Terminal	E-UTRAN	EPC	Mobile IMS	(Inter-national) trans-mission network	Far end access network	Far end terminal
Voice quality (MOS-LQxSW)	3.8 (if HD voice + TrFO) 2.8 (otherwise) Note 1	No target	No target	No target	No target	No target	No target	No target
Mouth-to-ear delay	400 ms Note 2	190 ms (sending + receiving) Note 3	80 ms	50 ms	0	see typical values in [b-GSMA IR.34] Table 7	UTRAN : 90	0
Call drop rate	3 %	No target	No target	No target	No target	No target	No target	No target

NOTE 1 – Assumes the use of AMR-WB at a 12.65 kbit rate. The use of other codecs and/or bit rates will result in other values.

NOTE 2 – [ITU-T G.114] specifies a preferred maximum value at 150 ms, impossible to reach currently; some network operators are able to provide national calls with delays below 300 ms.

NOTE 3 – According to [3GPP TS 26.131].

LTE-PSTN communication

Table 6 – Quality budgets for LTE-PSTN communication

End-to-end indicators	TOTAL budget	Terminal	E-UTRAN	EPC	Mobile IMS	(International) transmission network	PSTN
Registration success rate	99.9%	99.9%	99.9%	100%	99.9%	No target	No target
Service availability	99%	No target	No target	No target	No target	No target	100 %
Post dialling delay (PDD)	4 s CSFB: for further study	No target	No target	No target	No target	No target	No target
Voice quality (MOS-LQxSW)	3.1 (Note 1)	No target	No target	No target	No target	No target	No degradation
Mouth-to-ear delay	400 ms (Note 2)	190 ms (sending + receiving) (Note 3)	80 ms	50 ms	0	see typical values in [b-GSMA IR.34] Table 7	10 ms

Table 6 – Quality budgets for LTE-PSTN communication

End-to-end indicators	TOTAL budget	Terminal	E-UTRAN	EPC	Mobile IMS	(International) transmission network	PSTN
Call drop rate	2%	No target	No target	No target	No target	No target	0 %
NOTE 1 – Assumes the use of AMR at a 12.2 kbit rate. The use of other codecs and/or bit rates will result in other values. NOTE 2 – [ITU-T G.114] specifies a preferred maximum value at 150 ms, impossible to reach currently; some network operators are able to provide national calls with delays below 250 ms. NOTE 3 – According to [3GPP TS 26.131].							

10 Recommendations for QoS monitoring and troubleshooting

10.1 Recommended QoS monitoring points

The measurement points are critical to locate the issues as economical and operational matters required to limit the number of probes, robots or other tools.

The objective is to ensure a good representation of the covered territory in statistical terms for reporting, monitoring and troubleshooting with associated measurements sources or tools (probes, robots, counters and call detail records – CDRs).

Three kinds of measurements points can be envisaged (see Figure 10 below):

– At end points, where end-users access and experience the service (here A and J).

– At interfaces where the user plan is accessible, in general at demarcation points between network trunks or transmission technologies: here B, D, E and G.

– At points of presence of serving elements concerning signalling plan: here points C, F, and H (but also E).

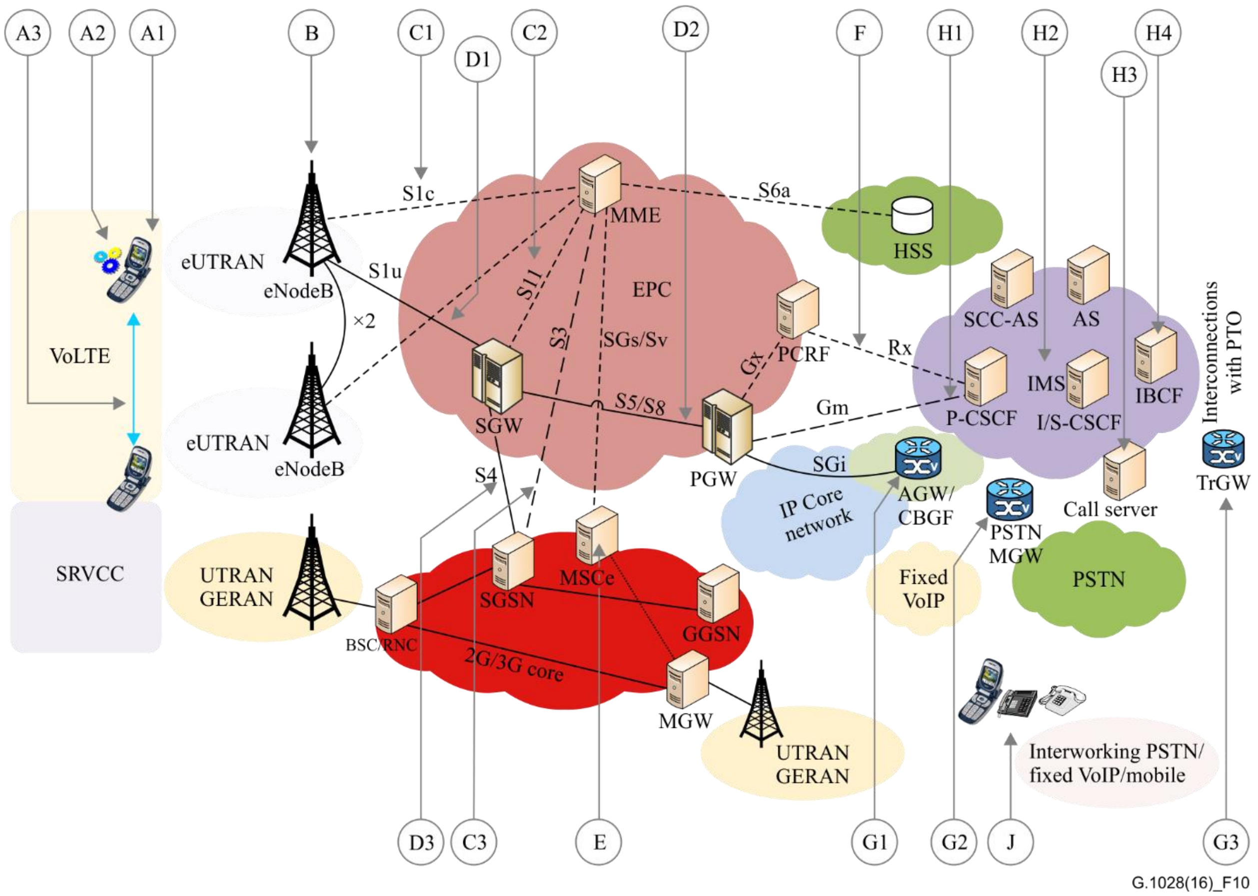

Figure 10 – VoLTE QoS measurement points

10.1.1 Measurement points A1, A2 and A3

Measurement at the customer premises is absolutely necessary to answer the important need of knowledge on how the service quality is felt by end users. It gives a real end-to-end visionas there is no other place to obtain it.

Then comes of course the question of representation: how many measurement points, what geographical repartition? The answer depends on the possible solutions at the concerned measurement points.

Different implementation are possible:

A1 measurement point is intended for measurements from a reporting point of view (manual test campaigns).

In fact, this point is not mandatory for supervision. But from an operator's point of view, a lot of QoS evaluations with this methodology are done each year on the entire network for regulatory or benchmark purposes. So, it is a way to check the QoS and to try to resolve encountered problems.

It enables to test end-to-end metrics on accuracy mostly, but also availability, and continuity. Such methodology is generally based on:

– A good and representative coverage of country area (rural, towns of different sizes, etc.).
– A selection of representative situation of communications (indoor, street, car, train, etc.).
– A selection of representative devices on the operator network.
– A selection of representative day periods to perform calls.
– A selection of representative calls. Narrowband and wideband calls must be taken into account.

– Enough measurements to have a correct view of QoS of the operator and relevant statistics.

A2 and **A3** measurement points are intended for measurements where tools replace the end user present in A1 and model his/her behaviour and quality judgment. They are recommended for reporting purposes, on end-to-end metrics on accuracy mostly, but also availability, continuity and service speed. This is where intrusive robots, drive test tools, radio tool analysers and soft agents are required. The cost of such solution stipulates that only a sampling strategy can be envisaged. The selection of measurement locations, devices and time periods follow similar criteria as for A1, but with less constraints since these are automatic tools.

A2 is inside the terminal. There soft agent can access decoded signal and/or information on how the communication is being processed (call set-up, call drop, IP jitter buffer behaviour, etc.). Such solutions are generally very cost-effective, can collect data on a very large panel of users and devices, but may require a participation or feedback of end-users.

A3 is where intrusive tools are required. This can be either static solutions based on robots (the focus will then be placed on service speed, accuracy and continuity) or mobile solutions (called drive test tools, with a larger focus enclosing also radio network performance). As seen before, for cost and operational reasons, only a sampling strategy can be envisaged with such tools (typically less than ten measurement points for national network). Moreover if a problem is reported by a measurement tool, it means that many users are probably impacted equally. As a consequence of the low number of measurements resulting from this sampling strategy, such tools will be mostly used for reporting purposes. Drive test tools are also commonly used for troubleshooting, when the enhancement of QoS is linked with the optimization of network coverage.

It is important to understand that this end-to-end vision enables the detection or the confirmation of a problem but rarely to know the source of the degradation, unless it comes from the device.

10.1.2 Measurement point B

The access part of the network is an important source of outages (biggest number of network elements). Therefore, its supervision is also necessary. One of the main difficulties is to take into account the radio part which can be clearly different according to the selected area.

The **B** point is localized in eNodeB, which is the first ingress point to the LTE network. This measurement point assumes monitoring the interfaces on both radio and in the EPC direction. The strategy assumes differentiation of VoIP traffic from other data traffic based on adequate traffic class indicator (QCI).

It is intended for measurements of:

– radio performance in cells: It is recommended notably for troubleshooting purposes, on metrics on network performance. Together with A2, the eNodeB point should enable a correct view of radio coverage performances of area;
– signalling flows, to establish statistics about calls: It is highly recommended for reporting, monitoring and troubleshooting purposes on end-to-end metrics on service availability, continuity, accuracy, utilization and network performance.

Supervision tools from RAN providers can enable a global aggregation of those kinds of information from all eNodeBs.

The measurement strategy for this point should be rather cheap in implementation (i.e., performance counters accessible from the device) than exhaustive.

10.1.3 Measurement points C1, C2 and C3

Points **C1** to **C3** corresponds to the interfaces of EPC (in fact all are localized close to mobility management entity (MME)) where most important signalling messages can be captured and analysed. They are recommended for reporting, monitoring and troubleshooting purposes on end-to-end metrics on service availability, continuity, and utilization. An exhaustive access to all data for all session is recommended, mostly through the usage of probes or network element call detail records (CDRs).

C1 covers signalling between MME and radio access (interface S1), **C2** supervises the link to serving GW (interface S11), whereas with **C3** the Sgs interface with MSC server is also addressed.

Please note that the traffic captured on all interfaces covered by C points is tunnelled (GTP v2-C).

These measurement point are important especially for the analysis of:

– EPS bearer statistics.

– Mobility management (HO statistics, routing area update).

– Subscriber management (direct connection to HSS).

10.1.4 Measurement points D1 and D2

Like C1, C2 and C3 these points are located inside the EPC, but covers interfaces where real time flows (using the RTP protocol) can be captured and analysed.

D1 is positioned at the egress between EPC and radio access (interface S1u of serving GW), whereas **D2** covers the link between serving and PDN GWs (interfaces S5-S8).

The serving GW constitutes the anchor point for intra-LTE mobility, as well as for mobility between GSM/GPRS, 3G/HSPA and LTE. It supports transport level QoS through marking IP packets with appropriate DiffServ code points based on the parameters associated with the corresponding bearer.

The PDN GW is the point of interconnect to external IP networks through the SGi interface. It also has a key role in supporting QoS for end-user IP services.

With such critical measurement points, all information on the quality of transport of communication is accessible. They are highly recommended for reporting, monitoring and troubleshooting purposes on end-to-end metrics on service availability, continuity, accuracy, speed and network performance. An exhaustive access to all data for all session is recommended, mostly through the usage of probes or network element CDRs.

10.1.5 Measurement point E

Measurement point **E** is located at the border between EPC and 3G core network. The corresponding equipment is the MSC server, which controls calls and mobility in 2G and 3G networks (e.g., in case of SRVCC). The MSC server is also in charge of TFO and TrFO functions. In many cases, MSC server is collocated with media gateway controller function (MGCF), in charge of creating and releasing connections at MGWs between IMS and CS domains.

This point is therefore recommended for monitoring and troubleshooting purposes on most metrics concerning communications where roaming between LTE and 2G or 3G occurs. The concerned metrics fall in the service availability, continuity, and speed families mostly.

10.1.6 Measurement point F

The policy and charging rule function (PCRF) is the central entity in policy control making policy and charging control (PCC) decisions. The decisions can be based on input from a number of different sources, including:

– Operator configuration in the PCRF that defines the policies applied to given services.

– Subscription information for a given user, received from the subscriber profile repository (SPR).

– Information about the service received from the application function (AF).

– Information from the access network about what access technology is used, etc.

The measurement point **F** is meant for monitoring the Rx interface between PCRF and the IMS domain proxy call session control function (P-CSCF). It is recommended for reporting, monitoring and troubleshooting purposes on metrics on service utilisation, availability, continuity, and speed.

10.1.7 Measurement points G1, G2 and G3

This group of measurement points corresponds to media gateways (GWs). GWs are important locations to have specific information about evaluation of core network status (IP packets losses, delay and jitter), codec in core, transcoding and use of TrFO. These are points of interest for monitoring and troubleshooting purposes on metrics such as service accuracy and network performance.

Proper monitoring of media flows (RTP) not only for interworking of services between EPC and other networks is supported by various types of GW:

– **G1** – IMS MGW – this is a GW used for anchoring media legs between two VoLTE users but also in case of SRVCC (HO from VoLTE to mobile CS) or VoLTE – legacy mobile CS interworking. This gateway should also provide functionality for interworking between VoLTE and fixed VoIP users (with AMR/AMR-WB – ITU-T G.711 transcoding).

– **G2** – PSTN MGW – this is a place where monitoring, reporting and troubleshooting can be applied (for network performance), for outgoing IP (from IP to PSTN) flows. This is also (and this is its most interest) the place where the performance of cancellation of echoes coming from PSTN is performed.

– **G3** – Trunking gateway (TrGW) in case of interconnection with another IP domain operator.

10.1.8 Measurement Points H1, H2, H3 and H4

This is the group of measurement points inside IMS platforms. Monitoring information which can be possessed here are related to the signalling traffic. Probes which can be deployed at the platform side should be complementary to the information received from CDRs and counters. Monitoring information which should be gathered at the corresponding measurement points are related to service availability, continuity and utilization. According to the service (voicemail, conferencing, etc.), accuracy and network performance supervisions can also be interesting to track audio problems.

Instead of having one single point, it is recommended to separate the capture of metrics at the level of different entities:

– **H1** – is located at the interface between EPC (PDN GW) and IMS (P-CSCF). This point covers passive monitoring of Gm interface, which is the first interface where none tunnelled SIP signalization coming from all users in the network is centralized;

– **H2** – is a global monitoring point localized on IMS platform (comprising at least elements such as I- and S-CSCF, AS and SCC-AS), ideal to make reporting on service utilization as the analysis will be highly statistically reliable. Monitoring data gathered in this point are focused on performance counters and billing data. As the most suitable platform element for gathering statistics may be metric dependent, detailed information on collection interface will be specified in metrics definitions;

– **H3** – measurement point (on call server/MGCF) can be used for an exhaustive view on the signalling protocol concerning all calls outgoing to PSTN network, since the call server is involved in all call negotiations. Service availability metrics (and other with less priority concerning continuity and network performance) are measurable at this point;

– **H4** – controls IBCF, where the interconnection between two operator domains is performed (including charging data records). SIP signalling information for interconnection can be monitored here.

10.1.9 Measurement point J

The **J** measurement point is the circuit-switched network counterpart of A point. The same metrics can be measured for similar purposes (except radio network performance), and with the same sampling strategy.

"Circuit switch", in this case means all possible interfaces: PSTN, 2G/3G PLMN as well as fixed VoIP access behind an FXS port of a DSL router.

10.2 Monitoring strategy

This clause describes the location of the metrics that have to be measured according to their usage. In particular, it aims at defining:

– what needs to be measured, how it will be measured (with what source of information, at what frequency) and why these measurements are needed;

– what kind of measurement tools and where (locations description and number) these measurements are needed.

10.2.1 Reporting

The reporting will provide a general view on the service QoS, potentially in a benchmark perspective (between countries, or against domestic competitors), and on its evolution in time (longitudinal view). It can also allow identification of categories of customers that have a degraded QoS.

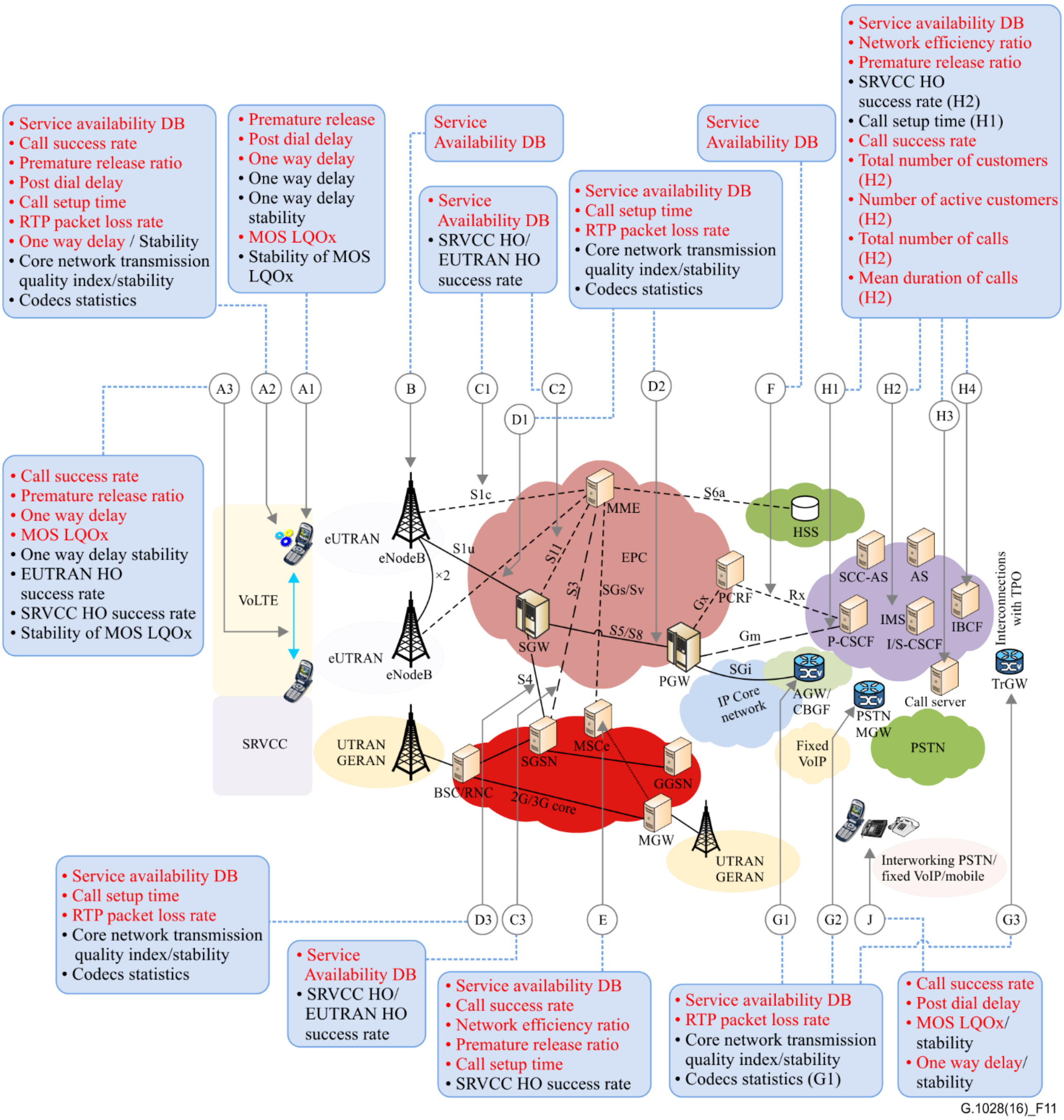

Figure 11 – VoLTE QoS measurement points – reporting

All measurement points can actually provide data for reporting. Even not fully representative data, like those gathered from intrusive measurements, can be valuable for this purpose.

The following dashboards can be built based on such measurements:

– General view of the service utilisation (number of customers, number of calls, call durations, churn rate).

– Performance of service platforms and network equipment (service availability and continuity).

– QoS counters (availability, PDD, mean opinion score (MOS), call continuity).

10.2.2 Monitoring

All measurement points can provide data for monitoring, as long as they can provide a view on either end-to-end QoS or the contribution of a network section to this QoS.

The hot mode monitoring (alarm raising) will provide real time information related to degradation of the service QoS impacting a great number of customers.

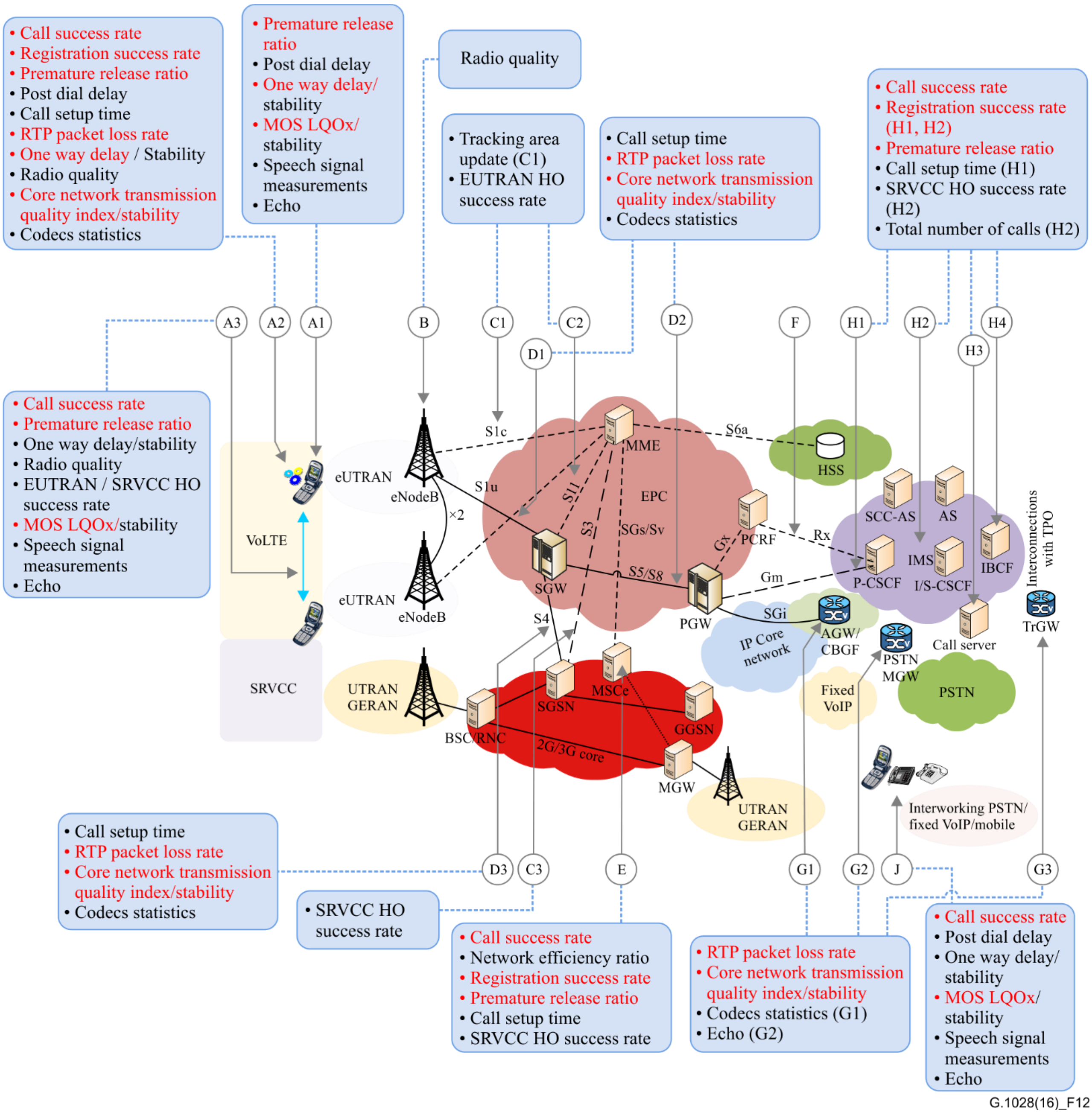

Figure 12 – VoLTE QoS measurement points – monitoring

10.2.3 Troubleshooting

The troubleshooting will provide additional information by:

– In depth analysis of customer or access below the threshold, by trying to reproduce the causes for the encountered problems and analysing data collected during the problem by different sources (CDRs, probes).

– Remote testing/configuration checks, including trace capture and analysis at major network nodes.

– On site investigations, including trace capture at the customer premise or in the access network.

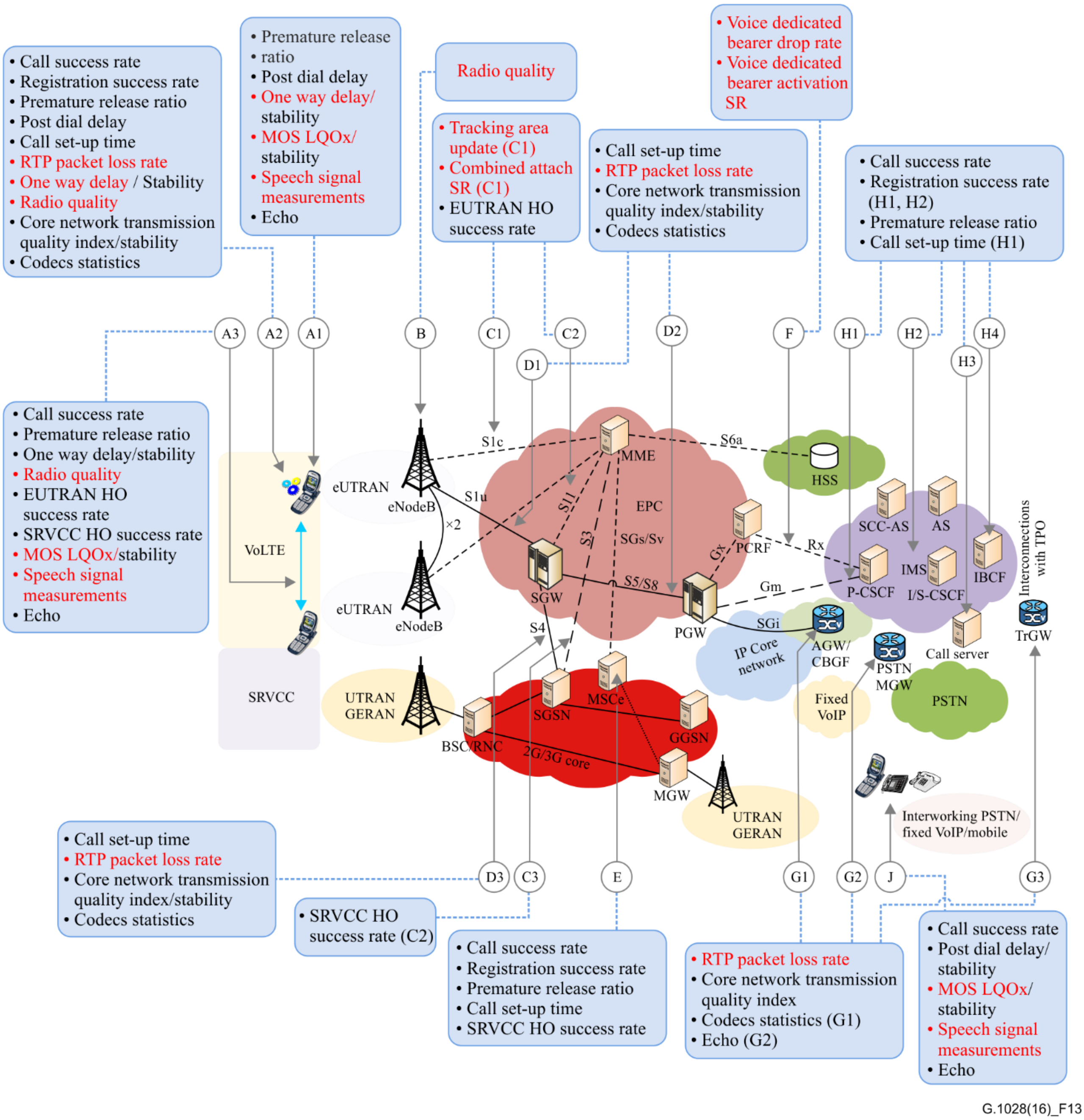

Figure 13 – VoLTE QoS measurement points – troubleshooting

10.3 Available tools in ITU-T standards

Though based on new network technologies, VoLTE remains a telephony offer subject to all requirements expressed by end users concerning its usage, and in particular when it comes to QoS. Therefore, provisions already contained in existing ITU-T Recommendations concerning the procurement and evaluation of perceived quality of voice services apply in most cases to VoLTE as well.

10.3.1 Definition of key performance indicators

As can be seen in previous sections (in particular in Table 2), the metrics representative of the quality of service as perceived by end users are not specific to VoLTE. Most of them are already defined in

[ITU-T P.10], and these definitions apply for this context. Following are the details of these references:

– Service availability: see service accessibility performance (clause 3.1.1.2.2 in [ITU-T E.800]).
– Post dialling delay: see call set-up time (clause 3.1.1.2.1 in [ITU-T E.800]).
– Voice quality (MOS-LQ, see "Speech Quality" (S-28) in [ITU-T P.10] or the various definitions provided in [ITU-T P.800.1].
– Mouth-to-ear delay: not defined as such in ITU-T Recommendations. See however mean one-way propagation time (M-1) in [ITU-T P.10], as well as explanation in [ITU-T G.114].
– Call drop ratio: see telephony cut off call ratio in [ITU-T E.804].
– Speech bandwidth (NB, WB or SWB): amendment 4 of [ITU-T P.10] gives very detailed definition of various audio bandwidth used in telephony.

Registration success rate for VoLTE is not defined in Recommendations. [ITU-T E.804] (clause 7.2.2.1) defines a network selection and registration failure ratio, but this disregards the specificity of IMS registration.

As far as the main QoS dimensions are concerned, the four viewpoints of QoS (see Figure 14) defined in [ITU-T G.1000] are also useful in the context of VoLTE.

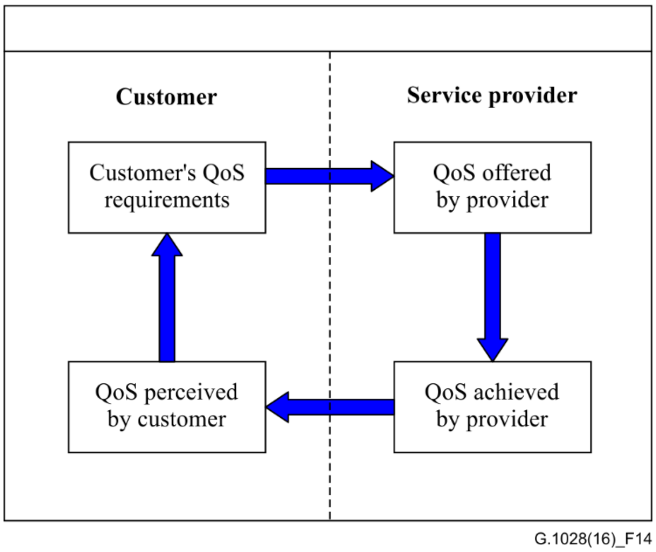

Figure 14 – The four viewpoints of QoS (source: [ITU-T G.1000])

As far as it concerns metrics representative of the underlying network layers, [ITU-T Y.1540] provides information on IP-related metrics, while no ITU-T standards address radio metrics.

10.3.2 Tools and models for the measurement and prediction of voice quality

There are two approaches for the assessment of end-to-end voice quality:

– Parametric tools take advantage of the good correlation between technical information of a connection and the corresponding end-to-end quality as perceived by end-users, to produce a relatively accurate estimate at a cheap implementation cost. Such a tool can be envisaged at edge points, close to the end-user, for a better prediction of individual quality, or inside the network, for a good knowledge of the general impact of network performance on end-to-end quality. [ITU-T P.564] describes a general class of parametric voice quality prediction models that provide highly scalable voice quality estimation using information in the IP/UDP/RTP header of packets. In addition, [ITU-T P.564] provides performance criteria for models of this type that operate on narrowband speech.

– Another type of parametric tool is the E-model described in [ITU-T G.107], which is a widely used transmission planning tool. Most factors present in the model (now adapted for IP transmission and WB telephony, see [ITU-T G.107.1]) apply for VoLTE.

– Signal-based models are much more complicated than parametric tools since they require a capture and an analysis of the speech signal. This is why they are mostly used in a point-to-point perspective, in order to measure very accurately the end-to-end quality of a voice service at a given time and at a given location (and most of the time for a given service with a given end device). ITU-T developed several such tools: [ITU-T P.862] and [ITU-T P.863] for full-reference models, and [ITU-T P.563] for a single-ended implementation (restricted to narrowband telephony).

10.3.3 Applicable acceptability thresholds and targets

In general, ITU-T do not specify acceptability targets for end-user's metrics. However, there is a notorious exception for end-to-end delay, where [ITU-T G.114] specifies (in clause 4) a first threshold at 150 ms below which most users do not notice the delay, and a second one at 400 ms above which the quality is considered as unacceptable.

R factor values computed with the E-model from [ITU-T G.107] (and translated into MOS-CQE scores, as shown on Figure 15) can also be compared with acceptability thresholds. Recommendation [ITU-T G.109] defines such categories (for NB telephony only) reproduced in Table 7:

Table 7 – Definition of categories of speech transmission quality (source: [ITU-T G.109])

R-value range	Speech transmission quality category	User satisfaction
$90 \leq R < 100$	Best	Very satisfied
$80 \leq R < 90$	High	Satisfied
$70 \leq R < 80$	Medium	Some users dissatisfied
$60 \leq R < 70$	Low	Many users dissatisfied
$50 \leq R < 60$	Poor	Nearly all users dissatisfied

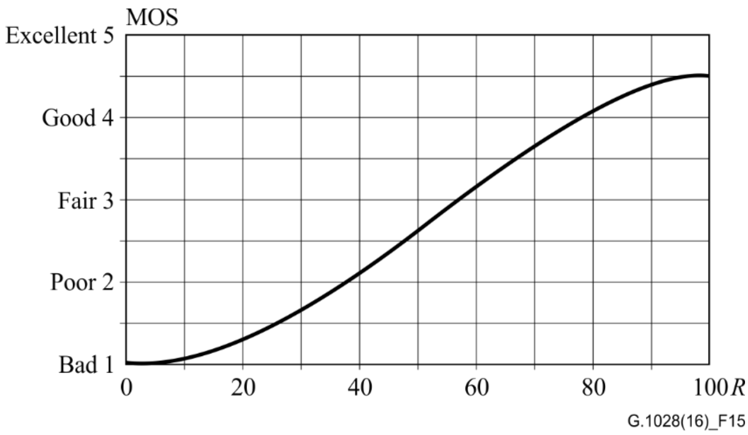

Figure 15 – MOS-CQE as function of rating factor (source: [ITU-T G.107])

Recommendation [ITU-T Y.1541] also provides performance objectives for various QoS classes (see Table 8) based on IP-network metrics defined and specified in [ITU-T Y.1540]. VoLTE falls into class 0 or 1:

Table 8 – IP network QoS class definitions and network performance objectives (source: [ITU-T Y.1541])

Network performance parameter	Nature of network performance objective	QoS classes	
		Class 0	**Class 1**
IPTD	Upper bound on the mean IPTD	100 ms	400 ms
IPDV	Upper bound on the $1 - 10^{-3}$ quantile of IPTD minus the minimum IPTD	50 ms	50 ms
IPLR	Upper bound on the packet loss probability	1×10^{-3}	1×10^{-3}
IPER	Upper bound	1×10^{-4}	

Annex A

List of degradations encountered by end-users of VoLTE service and their potential causes

(This annex forms an integral part of this Recommendation.)

A.1 QoS problem linked to call session performance

Table A.1 – degradations related to call session performance and their potential causes

<table>
<tr><th>Kind of degradation</th><th>Possible reasons:</th><th>Location</th></tr>
<tr><td>Identification Failure</td><td>– Problem with MME, HSS or PCRF</td><td>EPC</td></tr>
<tr><td rowspan="2">Unavailability of basic call</td><td>– Error in scheduling
– Radio resource control (RRC) connection set-up failure (reception of RRC connection reject, or expiry of timer T300, no RRC connection set-up complete sent after reception of RRC connection set-up).</td><td>eUTRAN</td></tr>
<tr><td>– Not available due to load (serving gateway (SGW) or packet data network gateway (PGW))
– Failed negotiation (no allocation of QCI, no codec match, SIP preconditions unmet, etc.)
– Reception of several SIP error codes (e.g., 401 = Unauthorized, 405 = Method Not Allowed, etc.)
– Reception of SIP CANCEL from IMS
– TD internal timer expired, causing a "SessionSetupFailureTimeout"</td><td>EPC</td></tr>
<tr><td>High post dialling delay</td><td>– Load
– Interworking between systems
– Use of SIP preconditions
– CS fall back or IMS tromboning at call set-up</td><td>All</td></tr>
<tr><td>Link failure</td><td>– Bad negotiation between two equipments of the network during call establishment (bad codec management)</td><td>eUTRAN/ EPC</td></tr>
<tr><td>White call</td><td>– Terminal is not able to code or decode speech while the signalling is OK for the communication</td><td>Terminal</td></tr>
<tr><td rowspan="2">Call drop</td><td>– Terminal bug, bad covered area, handover/SRVCC failures due to cells neighbourhood problem, etc.
– RRC connection drop (at reception of RRC connection re-establishment reject, or expiry of timer T301 or in case RRC connection release is received before new RRC connection set-up attempt)</td><td>Terminal/ eUTRAN</td></tr>
<tr><td>– Link failure: System failure, bad re-negotiation between two equipments of the network during call.
– Reception of SIP status code 500 (Server Internal Error)
– No RTP packet received during a period longer than "SessionDropTimeout" TD internal timer
– No SIP 200 OK on BYE is received within the time measured by "SessionHangupTimeout" TD internal timer</td><td>EPC</td></tr>
</table>

A.2 QoS problem linked to perceived speech quality

Table A.2 – Degradations related to perceived speech quality and their potential causes

Kind of degradation	Possible reasons:	Location
Noise	– Disturbing comfort noise generation (CNG) due to bad noise reduction.	Terminal
	– Noise due to bad electronic implementation on terminal (e.g., analogue /digital conversion).	
	– Disturbing residual noise due to bad noise reduction.	
	– Background noise (street, car, etc.).	
	– Additional noise due to eUTRAN configuration problem.	
Echo	– Bad performance of acoustic echo cancellation (AEC)/ No AEC. As reminder: Acoustic echo is the coupling between the loudspeaker and the microphone of the phone terminal.	
	– Bad performance of electric echo cancellation (EEC)/No EEC . Reminder: Electrical echo is due to digital to analogue transformation for a call between mobile terminal and PSTN (No electrical echo for mobile to mobile call).	Networks
Low/high speech level	– Bad performance of automatic gain control (AGC)/No AGC.	Terminal
Encoding / decoding issues	– Narrowband instead of wideband speech quality: • Remote terminal not WB • HO towards 2G • Call with PSTN, 2G, mobile platforms, etc. where wideband is not deployed. • Interworking with CS 3G not WB	Terminal/ eUTRAN
	– Lower WB-AMR bitrate/AMR (Loaded cell, autonomous mode, etc.) leading to distortion on speech signal. – Many transcodings (for example with call to voicemail) leading to distortion on speech signal – Rebuffering and time scaling causing distortion	Terminal/ eUTRAN
Terminal Acoustic	– Although WB-AMR codec is supported, the acoustical performance of the terminal (on receiving and/or sending side) is not wideband compliant. – Not well-balanced acoustic terminal can lead to a sound which seems too aggressive, too muffled, etc. – Distortion due to transducers.	Terminal
Chopped Conversation	– Bad VAD/DTX/DRX implementation. – Problem with voice quality enhancement (VQE) algorithm.	Terminal
	– IP packet loss or jitter in network (congestion, QoS prioritization, UL/DL scheduling delays, radio retransmissions, handover). – Bad handling of IP packet loss and inter-arrival jitter by jitter buffers or packet loss concealment (PLC) inside terminals	All
DTMF not recognized	– Problem with in-band or out-band processing	All

Table A.2 – Degradations related to perceived speech quality and their potential causes

Kind of degradation	Possible reasons:	Location
E2E delay	– Network load.– Media handling (packet construction, jitter buffer management) – Speech processing in terminals – Random access channel (RACH) upon receiving handover command – RACH/contention procedure – Additional RACH attempts – Dynamic scheduling, link adaptation – Radio link failure/re-establishment during handover (possibly different cell)	All
RTP/IP packet loss	– Network congestion (several causes : traffic load, distance from cell centre causing activation of TTI bundling, for instance) – Jitter buffers not adapted to actual jitter amount or packet size (can depend on use of RoHC or not)	EPC/ Terminal
RTP/IP Desequencing	– New route after a problem such as congestion	EPC
Network Delay Variation (Jitter)	– Network congestion – Jitter buffers not adapted	EPC/ Terminal
Radio degradations	– Limit of the cell coverage – Interference – Area not well covered (obstacle, etc.) – Bad radio optimization – Radio loss profile – Bad radio scheduling – No or bad use of hybrid automatic-repeat-request (HARQ) mechanisms – Etc.	eUTRAN
Handover latency	– Latency due to new route after HO or SRVCC	EPC/CS network

Bibliography

[b-GSMA IR.34] GSMA IR.34 v 9.1 (2013), *Guidelines for IPX Provider networks*.

[b-GSMA IR.92] GSMA IR.92 v 7.0 (2013), *IMS Profile for Voice and SMS*.

SERIES OF ITU-T RECOMMENDATIONS

Series A	Organization of the work of ITU-T
Series D	General tariff principles
Series E	Overall network operation, telephone service, service operation and human factors
Series F	Non-telephone telecommunication services
Series G	**Transmission systems and media, digital systems and networks**
Series H	Audiovisual and multimedia systems
Series I	Integrated services digital network
Series J	Cable networks and transmission of television, sound programme and other multimedia signals
Series K	Protection against interference
Series L	Environment and ICTs, climate change, e-waste, energy efficiency; construction, installation and protection of cables and other elements of outside plant
Series M	Telecommunication management, including TMN and network maintenance
Series N	Maintenance: international sound programme and television transmission circuits
Series O	Specifications of measuring equipment
Series P	Terminals and subjective and objective assessment methods
Series Q	Switching and signalling
Series R	Telegraph transmission
Series S	Telegraph services terminal equipment
Series T	Terminals for telematic services
Series U	Telegraph switching
Series V	Data communication over the telephone network
Series X	Data networks, open system communications and security
Series Y	Global information infrastructure, Internet protocol aspects and next-generation networks, Internet of Things and smart cities
Series Z	Languages and general software aspects for telecommunication systems

Printed in Switzerland
Geneva, 2016