



**5G;  
Study on media handling aspects of conversational services in  
5G systems  
(3GPP TR 26.919 version 16.2.0 Release 16)**



---

Reference

---

DTR/TSGS-0426919vg20

---

Keywords

---

5G**ETSI**

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

---

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

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

---

**Important notice**

---

The present document can be downloaded from:

<http://www.etsi.org/standards-search>

The present document may be made available in electronic versions and/or in print. The content of any electronic and/or print versions of the present document shall not be modified without the prior written authorization of ETSI. In case of any existing or perceived difference in contents between such versions and/or in print, the prevailing version of an ETSI deliverable is the one made publicly available in PDF format at [www.etsi.org/deliver](http://www.etsi.org/deliver).

Users of the present document should be aware that the document may be subject to revision or change of status.

Information on the current status of this and other ETSI documents is available at

<https://portal.etsi.org/TB/ETSIDeliverableStatus.aspx>

If you find errors in the present document, please send your comment to one of the following services:

<https://portal.etsi.org/People/CommiteeSupportStaff.aspx>

---

**Copyright Notification**

---

No part may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm except as authorized by written permission of ETSI.

The content of the PDF version shall not be modified without the written authorization of ETSI.

The copyright and the foregoing restriction extend to reproduction in all media.

© ETSI 2020.

All rights reserved.

**DECT™**, **PLUGTESTS™**, **UMTS™** and the ETSI logo are trademarks of ETSI registered for the benefit of its Members.

**3GPP™** and **LTE™** are trademarks of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners.

**oneM2M™** logo is a trademark of ETSI registered for the benefit of its Members and of the oneM2M Partners.

**GSM®** and the GSM logo are trademarks registered and owned by the GSM Association.

---

# 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.

---

# Legal Notice

This Technical Report (TR) has been produced by ETSI 3rd Generation Partnership Project (3GPP).

The present document may refer to technical specifications or reports using their 3GPP identities. These shall be interpreted as being references to the corresponding ETSI deliverables.

The cross reference between 3GPP and ETSI identities can be found under <http://webapp.etsi.org/key/queryform.asp>.

---

# Modal verbs terminology

In the present document "**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.

# Contents

Intellectual Property Rights .....	2
Legal Notice .....	2
Modal verbs terminology.....	2
Foreword.....	5
Introduction .....	5
1 Scope .....	6
2 References .....	6
3 Definitions and Abbreviations.....	7
3.1 Definitions .....	7
3.2 Abbreviations .....	8
4 5G System Overview.....	8
4.1 Introduction .....	8
4.2 Stage-1 Requirements.....	8
4.3 Stage-2 System Architecture .....	9
5 Impacts of 5G Stage-1 Requirements on 3GPP Conversational Services.....	11
5.1 Technical Aspect 1: Efficient Speech User Plane .....	11
5.1.1 Description.....	11
5.1.2 Implications on MTSI.....	11
5.1.3 Implications on IMS-based Telepresence .....	11
5.1.4 Recommended Requirements .....	11
5.1.5 Gap Analysis.....	11
5.1.6 Potential Solutions .....	12
5.2 Technical Aspect 2: Efficient Video User Plane .....	12
5.2.1 Description.....	12
5.2.2 Implications on MTSI.....	12
5.2.3 Implications on IMS-based Telepresence .....	12
5.2.4 Recommended Requirements .....	12
5.2.5 Gap Analysis.....	12
5.2.6 Potential Solutions .....	12
5.3 Technical Aspect 3: Media Rate Adaptation .....	12
5.3.1 Description.....	12
5.3.2 Implications on MTSI and IMS-based Telepresence.....	13
5.3.3 Recommended Requirements .....	13
5.3.4 Gap Analysis.....	13
5.3.5 Potential Solutions .....	13
5.4 Technical Aspect 4: VR Services .....	14
5.4.1 Description.....	14
5.4.2 Implications on MTSI.....	14
5.4.3 Implications on IMS-based Telepresence .....	14
5.4.4 Recommended Requirements .....	14
5.4.5 Gap Analysis.....	14
5.4.6 Potential Solutions .....	14
5.5 Technical Aspect 5: 5G New Radio (NR) Access.....	16
5.5.1 Description.....	16
5.5.2 Implications on MTSI.....	16
5.5.3 Implications on IMS-based Telepresence .....	17
5.5.4 Recommended Requirements .....	17
5.5.5 Gap Analysis.....	17
5.5.5.1 Overview of Gaps in MTSI.....	17
5.5.5.2 Setting MBR>GBR in 5GS .....	17
5.5.5.3 UE Adaptation at Rates > GBR .....	18
5.5.6 Potential Solutions .....	19

5.5.6.1	SDP Parameter for End-to-end RAN-assisted codec adaptation support .....	19
5.5.6.2	SDP Parameter for link-by-link RAN-assisted codec adaptation support .....	20
5.5.6.3	Support of RAN-Assisted Codec Adaptation on NR .....	20
5.6	Technical Aspect 6: Profiles for 5G Deployments .....	20
5.6.1	Description .....	20
5.6.2	Implications on MTSI .....	21
5.6.3	Implications on IMS-based Telepresence .....	21
5.6.4	Recommended Requirements .....	21
5.6.5	Gap Analysis .....	21
5.6.6	Potential Solutions .....	21
5.7	Technical Aspect 7: Support for Real-Time Interaction .....	21
5.7.1	Description .....	21
5.7.2	Implications on MTSI .....	22
5.7.3	Implications on IMS-based Telepresence .....	22
5.7.4	Recommended Requirements .....	22
5.7.5	Gap Analysis .....	23
5.7.6	Potential Solutions .....	23
5.7.6.1	WebRTC Data Channel Framework .....	23
5.7.6.1.1	Flexibility in Usage .....	23
5.7.6.1.2	Flexibility in Development .....	23
5.7.6.1.3	Flexibility in Operations .....	24
5.7.6.1.4	Relation to Existing WebRTC in IMS Specifications .....	24
5.7.6.2	MSRP .....	24
6	Impacts of 5G Stage-2 System Architecture on 3GPP Conversational Services .....	25
6.1	Technical Aspect 1: Mapping of Conversational Services to 5G System .....	25
6.1.1	Description .....	25
6.1.2	Implications on MTSI and IMS-based Telepresence .....	25
7	Codecs for 5G Conversational Services .....	25
7.1	Introduction .....	25
8	QoS Handling for 5G Conversational Services .....	25
8.1	Introduction .....	25
9	Conclusions .....	26
<b>Annex B:</b>	<b>Change history .....</b>	<b>27</b>
History .....		28

---

# Foreword

This Technical report has been produced by the 3<sup>rd</sup> Generation Partnership Project (3GPP).

The contents of the present document are subject to continuing work within the TSG and may change following formal TSG approval. Should the TSG modify the contents of the present document, it will be re-released by the TSG with an identifying change of release date and an increase in version number as follows:

Version x.y.z

where:

- x the first digit:
  - 1 presented to TSG for information;
  - 2 presented to TSG for approval;
  - 3 or greater indicates TSG approved document under change control.
- y the second digit is incremented for all changes of substance, i.e. technical enhancements, corrections, updates, etc.
- z the third digit is incremented when editorial only changes have been incorporated in the document.

---

# Introduction

The present document studies media handling aspects of 5G conversational services, focusing on Multimedia Telephony Service over IMS (MTSI) in TS 26.114 [4] and IMS-based Telepresence Service in TS 26.223 [5]. Various technical aspects including speech codes, video codecs, media rate adaptation, virtual reality (VR) support and new radio (NR) considerations are addressed, and related gap analysis and potential solutions are documented.

---

# 1 Scope

The present document provides a study on the media handling aspects of conversational services in 5G, taking as baseline the Stage-1 requirements developed in TS 22.261 [2], as well as the Stage-2 architecture for 5G systems developed in TS 23.501 [3]. This includes the investigation of the following areas:

- Media handling aspects of the 5G system architecture in relation to 3GPP conversational services, e.g., Multimedia Telephony Service over IMS (MTSI) in TS 26.114 [4] and IMS-based Telepresence Service in TS 26.223 [5].
- Relevance and potential reuse of components in existing 3GPP conversational services (e.g., MTSI, IMS-based telepresence, etc.) in the context of 5G systems and related Stage-2 architecture, e.g., use of MTSI features for supporting voice and video calls, use of MTSI, MS-MTSI and IMS-telepresence features for supporting multi-party conferencing, and applicability of existing QoE monitoring and QoS handling mechanisms.
- Potential enhancements to existing 3GPP conversational services (e.g., MTSI, IMS-based telepresence, etc.) towards better fulfilling the Stage-1 requirements in TS 22.261, e.g., in terms of criteria such as latency and bandwidth efficiency, while also taking into consideration the Stage-2 architecture for 5G systems:
  - In case existing codecs are unable to address 5G application requirements, new media codec requirements for 3GPP conversational services may be developed.
- The need for, and potential use of, new QoS media handling mechanisms in 5G systems such as traffic classification and codec-aware network elements in the context of 3GPP conversational services.

The gap analysis of the above areas and associated recommendations and conclusions for the proposed improvements are documented in the present document.

---

# 2 References

The following documents contain provisions which, through reference in this text, constitute provisions of the present document.

- References are either specific (identified by date of publication, edition number, version number, etc.) or non-specific.
- For a specific reference, subsequent revisions do not apply.
- For a non-specific reference, the latest version applies. In the case of a reference to a 3GPP document (including a GSM document), a non-specific reference implicitly refers to the latest version of that document *in the same Release as the present document*.

- [1] 3GPP TR 21.905: "Vocabulary for 3GPP Specifications".
- [2] 3GPP TS 22.261: "Service Requirements for Next Generation New Services and Markets; Stage 1".
- [3] 3GPP TS 23.501: "System Architecture for the 5G System; Stage 2".
- [4] 3GPP TS 26.114: "IP Multimedia Subsystem (IMS); Multimedia telephony; Media handling and interaction".
- [5] 3GPP TS 26.223: "Telepresence using the IP Multimedia Subsystem (IMS); Media handling and interaction".
- [6] 3GPP TR 23.799: "Study on Architecture for Next Generation System".
- [7] 3GPP TS 23.228: "IP Multimedia Subsystem (IMS); Stage 2".
- [8] 3GPP TS 23.503: "Policy and Charging Control Framework for the 5G System".
- [9] 3GPP TR 26.918: "Virtual Reality (VR) media services over 3GPP".

- [10] 3GPP TS 38.300: "NR; NR and NG-RAN Overall Description".
- [11] 3GPP TS 38.331: "NR; Radio Resource Control (RRC); Protocol Specification".
- [12] 3GPP TS 36.331: "Evolved Universal Terrestrial Radio Access (E-UTRA); Radio Resource Control (RRC) protocol specification".
- [13] 3GPP TR 26.910: "Study on Media Handling Aspects of RAN Delay Budget Reporting in MTSI".
- [14] 3GPP TR 26.959: "Study on enhanced Voice over LTE (VoLTE) performance".
- [15] ISO/IEC 23008-2: "Information technology -- High efficiency coding and media delivery in heterogeneous environments -- Part 2: High efficiency video coding".
- [16] ISO/IEC 23090-2: "Information technology -- Coded representation of immersive media -- Part 2: Omnidirectional media format".
- [17] 3GPP TS 26.118: "3GPP Virtual reality profiles for streaming applications".
- [18] IETF RFC 7798 (2016): "RTP Payload Format for High Efficiency Video Coding (HEVC)", Y.-K. Wang, Y. Sanchez, T. Schierl, S. Wenger, M. M. Hannuksela.
- [19] 3GPP TS 26.238: "Uplink Streaming".
- [20] 3GPP TS 38.306: "NR; User Equipment (UE) radio access capabilities".
- [21] 3GPP TS 38.321: "NR; Medium Access Control (MAC) protocol specification".
- [22] 3GPP TS 38.300: "NR; NR and NG-RAN Overall Description; Stage 2".
- [23] WebRTC 1.0: "Real-time Communication Between Browsers", W3C Candidate Recommendation, 27 September 2018, <https://www.w3.org/TR/2018/CR-webrtc-20180927/>
- [24] IETF Internet Draft draft-ietf-rtcweb-data-channel-13 (2015): "WebRTC Data Channel", <https://tools.ietf.org/html/draft-ietf-rtcweb-data-channel-13>, WORK IN PROGRESS.
- [25] IETF RFC 4566 (2006), "SDP: Session Description Protocol".
- [26] IETF Internet Draft draft-ietf-mmusic-data-channel-sdpneg-28 (2019): "SDP-based Data Channel Negotiation", <https://tools.ietf.org/html/draft-ietf-mmusic-data-channel-sdpneg-28>, WORK IN PROGRESS.
- [27] Node.js, Node.js Foundation, <https://nodejs.org/>
- [28] IETF RFC 4975 (2007), "The Message Session Relay Protocol (MRSP) "
- [29] 3GPP TS 24.247: "Messaging service using the IP Multimedia (IM) Core Network (CN) subsystem".
- [30] IETF Internet Draft draft-ietf-mmusic-msrp-usage-data-channel-12 (2019), "MSRP over Data Channels", <https://tools.ietf.org/html/draft-ietf-mmusic-msrp-usage-data-channel-12>, WORK IN PROGRESS.
- [31] 3GPP TS 24.371: "WebRTC access to the IMS; Stage 3; Protocol specification".

---

## 3 Definitions and Abbreviations

### 3.1 Definitions

For the purposes of the present document, the terms and definitions given in 3GPP TR 21.905 [1] and the following apply. A term defined in the present document takes precedence over the definition of the same term, if any, in 3GPP TR 21.905 [1].



## 3.2 Abbreviations

For the purposes of the present document, the abbreviations given in 3GPP TR 21.905 [1] and the following apply. An abbreviation defined in the present document takes precedence over the definition of the same abbreviation, if any, in 3GPP TR 21.905 [1].

5GC	5G Core
AF	Application Function
AMF	Access and Mobility Management Function
AN	Access Network
ANBR	Access Network Bitrate Recommendation
AS	Application Server
CMR	Codec Mode Request
CN	Core Network
CP	Control Plane
ECN	Explicit Congestion Notification
IBCF	Interconnection Border Control Function
KPI	Key Performance Indicator
MRFC	Media Resource Function Controller
MRFP	Media Resource Function Processor
MTSI	Multimedia Telephony Service over IMS
MS-MTSI	Multi-Stream MTSI
NEF	Network Exposure Function
NR	New Radio
NRF	Network Repository Function
OMAF	Omnidirectional Media Format
PCF	Policy Control Function
QoE	Quality of Experience
SA	Standalone
SEI	Supplemental Enhancement Information
SMF	Session Management Function
SRVCC	Single Radio Voice Call Continuity
TMMBN	Temporary Maximum Media Stream Bit Rate Notification
TMMBR	Temporary Maximum Media Stream Bit Rate Request
TrGW	Transition Gateway
UDM	Unified Data Management
UDR	Unified Data Repository
UL	Up-link
UP	User Plane
UPF	User Plane Function
VoLTE	Voice over LTE
VoNR	Voice over NR

---

## 4 5G System Overview

### 4.1 Introduction

This clause provides the background on 5G system requirements and architecture. Clause 4.2 describes the 5G Stage-1 Requirements. Clause 4.3 describes 5G Stage-2 architecture.

### 4.2 Stage-1 Requirements

TS 22.261 [2] developed by SA1 compiles service and operational requirements that define a 5G system. The 5G system is characterised, for example, by:

- Support for multiple access technologies
- Scalable and customizable network

- Advanced Key Performance Indicators (KPIs) (e.g., availability, latency, reliability, user experienced data rates, area traffic capacity)
- Flexibility and programmability (e.g., network slicing, diverse mobility management, Network Function Virtualization)
- Resource efficiency (both user plane and control plane)
- Seamless mobility in densely populated and heterogeneous environment
- Extreme long range coverage in low density areas
- Markets requiring minimal service levels with minimal user experience requirements around performance metrics such as bitrate, latency and coverage
- Support for real time and non-real time multimedia services and applications with advanced Quality of Experience (QoE), including services such as telepresence, virtual presence and telemedicine support

Many of the considered 5G system-related use cases and associated requirements in TS 22.261 relate to multimedia delivery, processing and storage.

## 4.3 Stage-2 System Architecture

In order to address the Stage-2 architectural aspects of 5G systems, SA2 has completed normative work "5G System - Phase 1", defined to support data connectivity and services enabling deployments to use techniques such as e.g. Network Function Virtualization and Software Defined Networking. For this purpose, the specification TS 23.501 [3] was developed, based on the conclusions of the Rel-14 study item FS\_NextGen and related TR 23.799 [6].

TS 23.501 covers the 5G System architecture that is defined to support data connectivity and services enabling deployments to use techniques such as Network Function Virtualization and Software Defined Networking. The 5G System architecture leverages service-based interactions between Control Plane (CP) Network Functions where identified. Some key principles and concept are to:

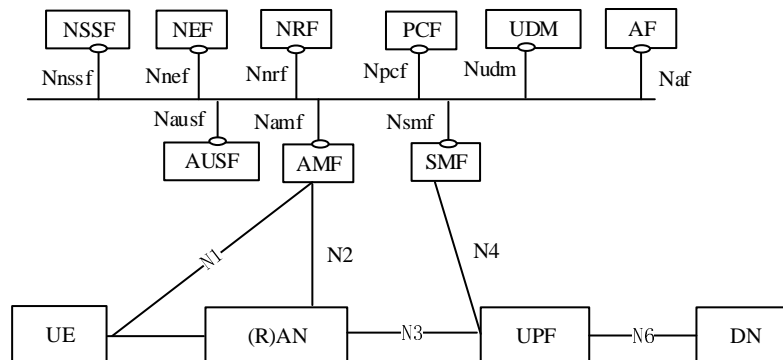
- Separate the User Plane (UP) functions from the Control Plane (CP) functions, allowing independent scalability, evolution and flexible deployments, e.g. at a centralized location or distributed (remote) locations.
- Modularize the function design, e.g. to enable flexible and efficient network slicing.
- Wherever applicable, define procedures (i.e. the set of interactions between network functions) as services, so that their re-use is possible.
- Enable each Network Function to interact with other NF directly if required. The architecture does not preclude the use of an intermediate function to help route Control Plane messages (e.g. like a DRA).
- Minimize dependencies between the Access Network (AN) and the Core Network (CN). The architecture is defined with a converged core network with a common AN - CN interface which integrates different 3GPP and non-3GPP access types.
- Support a unified authentication framework.
- Support "stateless" NFs, where the "compute" resource is decoupled from the "storage" resource.
- Support capability exposure.
- Support concurrent access to local and centralized services. To support low latency services and access to local data networks, UP functions can be deployed close to the Access Network.
- Support roaming with both Home routed traffic as well as Local breakout traffic in the visited PLMN.

The 5G architecture is defined as service-based and the interaction between network functions is represented in two ways.

- A service-based representation, where network functions (e.g. AMF) within the Control Plane enables other authorized network functions to access their services. This representation also includes point-to-point reference points where necessary.

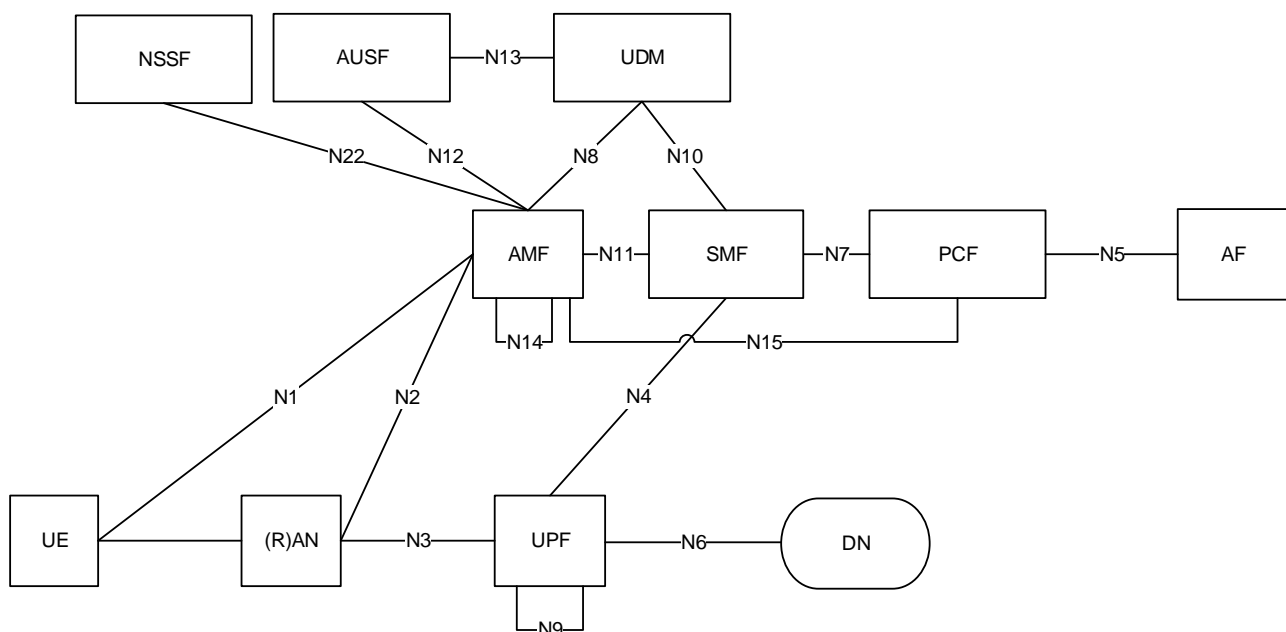
- A reference point representation, which shows the interaction between the NF services in the network functions described by point-to-point reference point (e.g. N11) between any two network functions (e.g. AMF and SMF).

Figure 4.3.1 depicts the non-roaming reference architecture. Service-based interfaces are used within the Control Plane.



**Figure 4.3.1: 5G System architecture**

Figure 4.3.2 depicts the 5G System architecture in the non-roaming case, using the reference point representation showing how various network functions interact with each other.



**Figure 4.3.2: Non-Roaming 5G System Architecture in reference point representation**

The 5G System architecture consists of the following network functions (NF):

- Application Function (AF) interacts with the 3GPP Core Network in order to provide services, for example to support the following functionalities: Application influence on traffic routing, accessing Network Exposure Function, interacting with the Policy framework for policy control.
- Access and Mobility Management function (AMF) includes the following functionalities: Mobility management, connection management, lawful intercept, transparent proxy, access authentication and authorization.
- Session Management Function (SMF) includes the following functionalities: Session establishment, modification and release, selection and control of UP function, UE IP address allocation and management, traffic steering configuration at UPF, control part of policy enforcement and QoS, charging data collection.

- User Plane Function (UPF) includes the following functionalities: Packet routing & forwarding, packet inspection, user plane part of policy rule enforcement, lawful intercept (UP collection), traffic usage reporting, external PDU session point of interconnect to data network.
- Policy Control Function (PCF) includes the following functionalities: Providing policy rules to control plane functions to enforce them, serving as a front end to access subscription information relevant for policy decisions.
- Network Exposure Function (NEF) includes the following functionality: Providing means to securely expose the services and capabilities of the 3GPP network to third parties, including application functions and edge computing.
- Network Repository Function (NRF) includes the following functionalities: Support of service discovery, maintaining NF profile of available NF instances and providing information of the discovered NF instances.
- Unified Data Management (UDM) includes the following functionalities: 3GPP AKA Authentication Credential Processing, User Identification Handling, Access Authorization, Registration/Mobility management, Subscription management, SMS management.
- Unified Data Repository (UDR) includes the following functionalities: Storage and retrieval of subscription data by the UDM, storage and retrieval of policy data by the PCF, storage and retrieval of application data (including packet flow descriptions) by the NEF.

---

## 5 Impacts of 5G Stage-1 Requirements on 3GPP Conversational Services

### 5.1 Technical Aspect 1: Efficient Speech User Plane

#### 5.1.1 Description

Clause 6.1 of TS 22.261 [2] requires an efficient 5G user plane. While the requirement text focuses on efficient user plane routing, user plane "efficiency" in the sense of achieved media quality per transmitted bit should also be in scope.

#### 5.1.2 Implications on MTSI

Current MTSI speech does not mandate use of the most bit-efficient and highest quality speech, but leaves it optional. To better meet 5G requirements, the most bit-efficient, highest quality speech should be made mandatory.

#### 5.1.3 Implications on IMS-based Telepresence

Current IMS-based Telepresence already mandates use of the most bit-efficient and highest quality speech, so there is no foreseen impact from this aspect.

#### 5.1.4 Recommended Requirements

It is recommended that 5G MTSI UE support for AMR and AMR-WB codecs is mandated, so as to minimize transcoding needs.

It is also recommended that the level of support for EVS and SWB operation be developed.

#### 5.1.5 Gap Analysis

In current MTSI, super-wideband speech support is optional, and the most bit-efficient wideband and narrowband codecs are optional.

### 5.1.6 Potential Solutions

A solution is to mandate for MTSI clients in terminals offering speech communication to support narrowband, wideband and super-wideband communication.

## 5.2 Technical Aspect 2: Efficient Video User Plane

### 5.2.1 Description

Clause 6.1 of TS 22.261 [2] requires an efficient 5G user plane. While the requirement text focuses on efficient user plane routing, user plane "efficiency" in the sense of achieved media quality per transmitted bit should also be in scope.

### 5.2.2 Implications on MTSI

Currently, it is not mandated for video-capable MTSI endpoints to use the most bit-efficient and highest quality video, but it is left as optional. To better meet 5G requirements, it should be mandatory for video-capable 5G MTSI endpoints to use the most bit-efficient, highest quality video.

### 5.2.3 Implications on IMS-based Telepresence

Currently, it is not mandated for video-capable IMS-based Telepresence endpoints to use the most bit-efficient and highest quality video, but it is left as optional. To better meet 5G requirements, it should be mandatory for video-capable 5G IMS-based Telepresence endpoints to use the most bit-efficient, highest quality video.

### 5.2.4 Recommended Requirements

It is recommended that support for H.265/HEVC (with the levels and profiles currently recommended respectively in TS 26.114 and TS 26.223) is mandated for video-capable 5G MTSI and IMS Telepresence clients in terminals.

In addition, it is recommended that 5G MTSI UEs support H.264 (AVC) constrained baseline profile level 1.2, and that 5G IMS Telepresence UEs support H.264 (AVC) constrained high profile level 3.1 and H.264 (AVC) constrained baseline profile level 1.2, as currently specified in TS 26.114 and TS 26.223, as to minimize transcoding needs.

### 5.2.5 Gap Analysis

Currently, H.265/HEVC support is optional for MTSI and IMS Telepresence endpoints.

### 5.2.6 Potential Solutions

A solution to achieve support for the most bitrate-efficient and highest quality video in 5G MTSI, is to mandate support in video-capable endpoints for the, currently optional, H.265/HEVC Main Profile, Main Tier, Level 3.1.

A solution to achieve support for the most bitrate-efficient and highest quality video in 5G Telepresence, is to mandate support in video-capable endpoints for the, currently optional, H.265/HEVC Main Profile, Main Tier, Level 4.1.

## 5.3 Technical Aspect 3: Media Rate Adaptation

### 5.3.1 Description

Clause 6.16 of TS 22.261 [2] requires handling markets requiring minimal service levels. Interoperability between such markets and other markets without such minimal service levels would be improved if the higher service level markets are capable to automatically adapt to the lower service level.

Clause 6.17 of TS 22.261 [2] requires handling extreme long range coverage in low density areas. It can be assumed that extreme long range coverage would also mean use of very low media bitrates, as well as dynamically removing or adding entire media components, when passing some minimum bitrate threshold for inclusion of that media component

in the session. Transcoding-free interoperability between UE camping in low density area and UE in other areas without such major bitrate limits would be improved if UE without bitrate limitations are capable to automatically and dynamically adapt to the UE having such limitation.

Clause 7.1 of TS 22.261 [2] requires handling high data rates and traffic densities. Seen from a media perspective for an individual UE, handling high data rates allows using high media bitrates. On the other hand, handling high traffic densities can mean that the same, individual UE gets allotted a fairly low media bitrate. Since traffic density as well as general radio conditions can be expected to vary over time, the available media bitrate can in general also be expected to vary over time, even during a single call. Meeting those requirements will be easier if the UE has the capability to automatically and dynamically adapt its sending bitrate to match available bitrate on the local uplink, and assist the remote media sender to match available bitrate on the local downlink. End-to-end transcoding-free operation is enabled by dynamically adapting sending bitrate to the minimum of the total end-to-end media path, including local uplink and remote downlink.

Clause 7.2 of TS 22.261 [2] requires handling low latency and high reliability. Seen from a media perspective for an individual UE, a key part of keeping low media latency and high reliability is to, at any point in time, avoid sending higher media bitrate than what the end-to-end media path can currently support. Such media bitrate is assumedly always lower than the maximum allowed, negotiated bandwidth, but there can also be more dynamically varying limitations that are lower than that negotiated upper limit, e.g. caused by other network traffic and/or various types of physical limitations in the end-to-end media path. Sending a higher bitrate than what the end-to-end media path can transport, can either cause data buffering, which increases media latency, or if buffering capabilities are limited, instead decrease reliability by causing excess media data loss/discard. Meeting those requirements will be easier if the UE has the capability to automatically and dynamically adapt its sending bitrate to match available bitrate on the end-to-end media path.

### 5.3.2 Implications on MTSI and IMS-based Telepresence

Current MTSI and IMS-based Telepresence does not mandate support and use of all bitrate adaptation functionality included in the respective specification, but leaves some parts of it optional. To better meet 5G requirements, the best possible bitrate adaptation using available methods should be made mandatory.

### 5.3.3 Recommended Requirements

It is recommended that best possible bitrate adaptation using available methods is made mandatory for 5G MTSI and IMS Telepresence.

### 5.3.4 Gap Analysis

In current MTSI and IMS Telepresence, use of a=bw-info, ANBR, ECN, sending CMR, and video triggers capable to detect 10% or more reduction in video bitrate are optional.

### 5.3.5 Potential Solutions

A tentative solution to enhance the bitrate adaptation with available methods in 5G MTSI and IMS Telepresence is to recommend the support for the following capabilities for MTSI and TP clients:

- If so configured by the operator, support of ANBR as an adaptation trigger, as described by clause 10.7 of TS 26.114 [2].
- If so configured by the operator, speech media receiver capability to trigger sending CMR requesting bitrate adaptation in the corresponding media sender RTP stream, based on estimated media receive direction channel quality, as described in clause 10.2 of TS 26.114 [2]. When the UE is triggering adaptation in this manner, an example adaptation logic may be provided.
- Video media receiver capability to use one or more adaptation triggers, jointly capable to detect a needed reduction in throughput of 10% or more, and to send RTCP TMMBR accordingly.
- Video media sender taking all of the above adaptation triggers into account and adapting the sent bitrate to the highest bitrate that is still lower than or equal to the minimum of the adaptation triggers, and to send RTCP TMMBN accordingly. When the UE is triggering adaptation in this manner, an example adaptation logic may be provided.

It should be possible for an operator to configure MTSI clients according to its policy on whether or not to use adaptation and offer these capabilities in the SDP. This could also be achieved by the network using B2BUA to disable use of adaptation by modifying the SDP.

## 5.4 Technical Aspect 4: VR Services

### 5.4.1 Description

Clause 7.2.3 of TS 22.261 [2] includes requirements toward supporting virtual reality (VR) and interactive conversation use cases, including relevant motion-to-photon and motion-to-sound latency requirements. Moreover, clauses 5.9, 5.10 and 5.11 of TR 26.918 [9] contain use cases on conversational VR (i.e., spherical video calls, videoconferencing with 360 video), user-generated VR live streaming (i.e., "See what I see") and virtual world communication, respectively, involving interactive real-time encoding, delivery and consumption of VR content relevant for MTSI and IMS-based telepresence.

### 5.4.2 Implications on MTSI

Currently, MTSI endpoints as specified in TS 26.114 [4] does not support mechanisms to encode, deliver and consume VR content.

### 5.4.3 Implications on IMS-based Telepresence

Currently, IMS-based Telepresence endpoints as specified in TS 26.223 [5] does not support mechanisms to encode, deliver and consume VR content.

### 5.4.4 Recommended Requirements

It is recommended that 5G MTSI and IMS Telepresence endpoints support VR capabilities relevant for real-time encoding, delivery and consumption of 3D spatial audio and 360 videos toward fulfilling the recommended objectives in clause 9.3.2 of TR 26.918 [9].

### 5.4.5 Gap Analysis

The gap analysis presented in clauses 5.9.3 and 9.3.1 of TR 26.918 [9] is applicable for 5G MTSI and IMS Telepresence endpoints.

On top of these gaps, the following detailed gaps specific to 5G MTSI and TP endpoints may be listed:

- To enable VR support, it is necessary to define SDP-based mechanisms for the negotiation of VR capabilities across MTSI / TP senders and receivers during both call setup and mid-call. VR capabilities here include the related codecs, formats and media handling mechanisms for encoding, delivery and consumption of 3D spatial audio and 360 degree videos. This for instance includes the negotiation of relevant projection and/or packing format(s) (including fish-eye video) to be used by the sender and receiver during delivery of 360 degree videos.
- In addition, for interactive navigation, it is beneficial to define suitable formats for real-time signalling of field-of-view (FOV) or viewport information from an MTSI / TP receiver to an MTSI / TP sender and vice versa during a multimedia telephony session, including consideration of the related RTP/RTCP based protocol impacts. Such indication of the viewport can allow for encoding optimizations on the sender side, toward delivering a higher quality stream and/or reduction of the bandwidth consumption.

### 5.4.6 Potential Solutions

Clause 9.3.3 of TR 26.918 [9] documents a few potential solutions.

Further potential solutions are for further study.

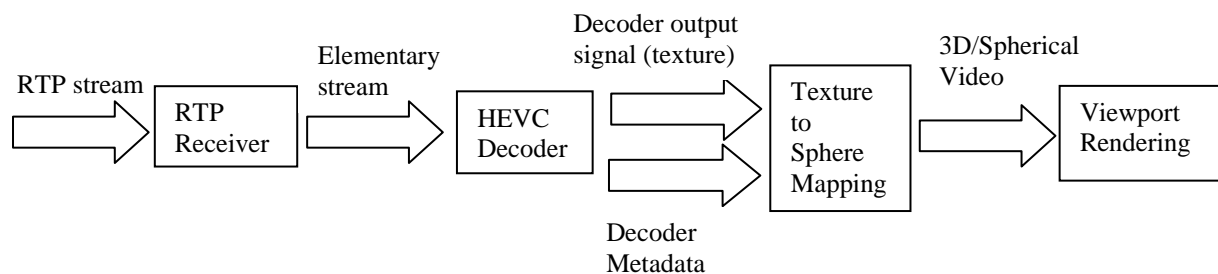
The relevant interoperability points for VR support over MTSI and IMS-telepresence are:

- Media profiles providing RTP and elementary stream constraints for a single media type.

- Rendering Scheme types for post-decoder processing of decoder output signals together with rendering metadata.
- Potential Viewport test points for rendered output signals.

Note that this applies to both media types, audio and video. The elementary stream constraints of a media profile may be indicated by a requirement to comply with a certain profile and level of the media coding specification, possibly including additional constraints and extensions, such as a requirement of the presence of certain information for rendering and presentation.

Figure 5.4.6.1 provides an overview of a possible receiver architecture that recovers the spherical video in an MTSI or IMS Telepresence UE. Note that this figure does not represent an actual implementation, but a logical set of receiver functions. Based on a received RTP media stream, the UE parses, possibly decrypts and moves the elementary stream to the HEVC decoder. The HEVC decoder obtains the decoder output signal, referred to as the "texture", as well as the decoder metadata. The Decoder Metadata contains the Supplemental Information Enhancement messages to be used in the rendering phase. In particular, the SEI messages may be used by the Texture-to-Sphere Mapping function to generate a spherical video based on the decoded output signal, i.e., the texture. The viewport is then generated from the spherical video signal by taking into account viewport position information from sensors, display characteristics as well as possibly other metadata such as initial viewport information.



**Figure 5.4.6.1: Potential receiver architecture for VR support over MTSI and IMS Telepresence**

In addition, the definitions and reference systems in clause 4.1 of TS 26.118 [17] are applicable to VR support over MTSI and IMS telepresence.

**NOTE:** Considering that real-time encoding of 360 degree video content could have significant latency, the initial consideration of VR support over MTSI is not on conversational VR, but rather on the user generated live VR streaming use cases. The use of MTSI in this context relies on live encoding of VR content, as in the case of Framework for Live Uplink Streaming (FLUS) in TS 26.238 [19].

For 360 degree video, the potential solutions can consider the following principles:

- The RTP stream would contain an HEVC bitstream with Supplemental Enhancement Information (SEI) messages in regards to omnidirectional media. In particular, SEI messages describing the decoder rendering metadata on the omnidirectional video as defined in ISO/IEC 23008-2 [15] may be present.
- Video elementary streams may be encoded following the requirements in the Omnidirectional Media Format (OMAF) specification ISO/IEC 23090-2 [16], clause 10.1.2.2.

Relevant SEI messages contained in the elementary stream with decoder rendering metadata may include the following information as per ISO/IEC 23008-2 [15]:

- Region-wise packing information, e.g., carrying packing format indication and also any coverage restrictions
- Projection mapping (indicates projection format), with indication of Equi-Rectangular projection (ERP) or Cubemap projection
- Padding, indicates whether there is padding in the packed frame
- Frame packing arrangement, indicating packing format for stereoscopic content
- Content prerotation information, including sphere rotation



The output signal, i.e. the decoded picture or "texture", is then rendered using the SEI messages contained in the video elementary streams. Metadata is used when performing rendering operations such as region-wise unpacking, projection de-mapping and rotation toward creating spherical content for each eye.

Viewport-dependent processing could be achieved by sending from the MTSI receiver RTCP feedback or RTP header extension messages with the desired viewport information and then encoding and sending the corresponding viewport by the MTSI sender. This is expected to deliver resolutions higher than the viewport independent approach for the desired viewport. For one-to-one video telephony scenarios, approaches such as tiling and sub-picture coding in the viewport-dependent profile of OMAF in ISO/IEC 23090-2 [16] etc. are not relevant for the 5G conversational setting, as the MTSI sender can customize the encoding according to the viewport chosen by the MTSI receiver and signalled to the MTSI sender using RTCP feedback or RTP header extension messages. However, viewport-dependent processing based on tiling and sub-picture coding could be relevant for multi-party video conferencing scenarios.

OMAF Video profiles in ISO/IEC 23090-2 [16] are based on HEVC Main 10 Profile, Main Tier, Level 5.1 in order to deliver high quality VR experiences. In the meantime, MTSI in TS 26.114 [4] recommends H.265 (HEVC) Main Profile, Main Tier, Level 3.1 for video, and IMS telepresence in TS 26.223 [5] recommends H.265 (HEVC) Main Profile, Main Tier, Level 4.1 for video.

For achieving video quality required by VR services, it may be recommended that the video codecs for VR support in MTSI and IMS telepresence are aligned with OMAF and/or TS 26.118 [17]. It is expected that both MTSI client and MTSI gateway codec requirements are aligned with these recommended video codec requirements for VR support. It is not expected that the mechanisms for session setup and negotiation would be different because of this changed requirement on video codecs.

With regards to the negotiation of SEI messages for carriage of decoder rendering metadata, procedures specified in IETF RFC 7798 [18] on the RTP payload format for HEVC may be reused. In particular, RFC 7798 can allow exposing SEI messages related to decoder rendering metadata for omnidirectional media in the SDP using the 'sprop-sei' parameter, which allows to convey one or more SEI messages that describe bitstream characteristics. When present, a decoder can rely on the bitstream characteristics that are described in the SEI messages for the entire duration of the session. Intentionally, RFC 7798 does not list an applicable or inapplicable SEI messages to be listed as part of this parameter, so the newly defined SEI messages for omnidirectional media in ISO/IEC 23008-2 [15] can be signalled. It is expected that both MTSI clients and MTSI gateways support RTP payload formats for VR support.

For one-to-one video telephony scenarios, it is expected that support of the following omnidirectional video specific SEI messages would be sufficient: 1) the equirectangular projection SEI message, 2) the cubemap projection SEI message, and 3) the sphere rotation SEI message. For multi-party video conferencing scenarios, in addition to the above three SEI messages, the support of the region-wise packing SEI message could also be useful. For stereoscopic video support, in either one-to-one video telephony scenarios or multi-party video conferencing scenarios, support of the frame packing arrangement SEI message is needed.

## 5.5 Technical Aspect 5: 5G New Radio (NR) Access

### 5.5.1 Description

5G New Radio (NR) has been defined by 3GPP as a new access technology [10] - [11] for Release 15. The 5G system will support 3GPP access technologies, including one or more NR and E-UTRA as well as non-3GPP access technologies. Standalone (SA) NR will be connected to 5GC and use 5G QoS. Interoperability among the various access technologies will be imperative. For optimization and resource efficiency, the 5G system will select the most appropriate 3GPP or non-3GPP access technology for a service, potentially allowing multiple access technologies to be used simultaneously for one or more services active on a UE. Clause 5.1 of TS 22.261 [2] includes high level requirements toward supporting 5G access technologies, including NR. Clause 6.3 of TS 22.261 contains requirements for interworking with the various combinations of 5G access technologies including NR.

### 5.5.2 Implications on MTSI

5G NR is a new type of access for an MTSI client to connect to IMS.

Currently, MTSI endpoints as specified in TS 26.114 [4] do not address media handling aspects with 5G NR access.

### 5.5.3 Implications on IMS-based Telepresence

5G NR is a new type of access for an IMS-based Telepresence client to connect to IMS.

Currently, IMS-based Telepresence endpoints as specified in TS 26.223 [5] do not address media handling aspects with 5G NR access.

### 5.5.4 Recommended Requirements

It is recommended that media handling for NR access is supported for 5G MTSI and IMS Telepresence endpoints, including that for speech and video.

NOTE: NR is not necessarily the preferred access over LTE access. It is TBD on what the preferred access would be when multiple access technologies are available.

### 5.5.5 Gap Analysis

#### 5.5.5.1 Overview of Gaps in MTSI

From MTSI perspective, the introduction of NR as a new type of access leads to several gaps in TS 26.114 [4], some of which can be listed as follows:

- In Tables 7.1 and 12.1 of TS 26.114 [4] for the MTSI client and MTSI media gateway, respectively, the speech frame encapsulation parameters for the SDP offer-answer messages including ptime and maxptime need to be defined for NR as a new radio access bearer technology.
- Access Network Bitrate Recommendation (ANBR) (as defined in clause 10.7 of TS 26.114 [4]) information may be signalled to the MTSI client in the UE using NR access, as defined in TS 38.300 [10], TS 38.331 [11], TS 38.306 [20] and TS 38.321 [21], and in this scenario the message mapping to NR access needs to be provided.
- Explicit Congestion Notification (ECN) in NR is specified in [22], and hence ECN-triggered media rate adaptation for speech and video is possible for MTSI clients with NR access. MTSI clients connected with NR access may optionally offer ECN, and relevant recommendations on use of ECN in MTSI that are already present in TS 26.114 [2] are also applicable for MTSI clients with NR access.
- RAN delay budget reporting mechanisms, as specified in TS 36.331 [12] for LTE access, may also be used by MTSI UEs with NR access in order to locally adjust air interface delay, towards improving end-to-end delay and quality performance. RAN delay budget reporting for NR access is specified in TS 38.300 [10], TS 38.331 [11] and TS 38.306 [20]. As such, media handling aspects of RAN-based delay budget reporting as described in TR 26.910 [13] are applicable for MTSI clients with NR access.
- Media handling enhancements for enabling codec-aware optimizations of Single radio voice call continuity (SRVCC) handover thresholds in Voice over LTE (VoLTE) as described in TR 26.959 [14] are also applicable for speech services to MTSI clients with NR access, including fallback/handover from Voice over NR (VoNR) to VoLTE.

NOTE: In the first release of NR for Rel-15, the interworking with 2G/3G is not supported, and VoNR can only be supported via IMS. SRVCC from 5GS to UTRAN support for VoNR is work in progress.

For IMS-based telepresence in TS 26.223 [5], the same gaps observed for MTSI above are applicable and no further gaps are foreseen.

#### 5.5.5.2 Setting MBR>GBR in 5GS

TR 23.860, clause 6.5, contains the following text:

*To make MBR>GBR bearers useful for a 3GPP application / service (e.g., MTSI) based on a bit rate adaptive codec requires that when sending beyond GBR the media end-points become aware of incipient congestion ahead of time. This is to allow the media end-point to trigger a codec rate reduction before packets need to be dropped in the network. With the Explicit Congestion Notification (ECN) scheme supported for UTRA/HSPA and E-UTRA, and for voice and video Codecs this requirement is met. So given that the media end-points have successfully negotiated the use of ECN no problem has been identified with simply allowing MBR>GBR bearers.*

*No additional functionality has been identified beyond what has already been specified in Rel-8 that would be required from a UE to support MBR>GBR bearers. On NAS level a UE shall anyway not reject a dedicated bearer based on QoS parameters (see dedicated establishment procedure in 23.401: "The UE may provide the EPS Bearer QoS parameters to the application handling the traffic flow. The application usage of the EPS Bearer QoS is implementation dependent. The UE shall not reject the RRC Connection Reconfiguration on the basis of the EPS Bearer QoS parameters contained in the Session Management Request.").*

*SA4 has stated [S4-070314] that for 3GPP services like MTSI packet dropping is not an acceptable means to trigger a codec rate reduction. In particular for video the error propagation may greatly damage video quality. SA WG4 even states: "Dropping packets may actually increase the media bit rate if e.g. frame redundancy is invoked due to degraded channel conditions." This is why an operator may want to configure a policy that the PCRF shall check that the end-points have successfully negotiated the use of ECN, i.e. that a congestion pre-warning scheme has been put in place, before the network sets MBR>GBR for 3GPP services like MTSI. Note that MBR>GBR is allowed already in 2G/3G pre-Rel-8.*

The RAN-assisted codec adaptation feature was developed in Rel-14 as the ECN-based solution was not reliable as it is difficult to guarantee that intermediate routers in the core network and along the transmission path would not drop ECN-marked packets. The ECN-based adaptation also had these other limitations:

- 1) It could not provide an explicit indication of what rate could be supported when congestion was experienced or eased. Adaptation could easily over-shoot or be too-slow, especially when ramping up rates after congestion eased.
- 2) It could not be reliably confirmed that ECN was supported by each access network in the transmission path. An eNB that was not ECN-aware may allow the end-to-end negotiation of ECN without it performing CE markings if it experienced congestion.

RAN-assisted codec adaptation has the advantages that:

- 1) It is implemented in the access network and UEs. It does not rely on changes to the core network or intermediate routers.
- 2) It provides an explicit indication of the UL and DL rate that is supported, enabling faster and more accurate adaptation.

RAN-assisted codec adaptation has the disadvantages that:

- It does not allow intermediate routers in the core network or backhaul to indicate congestion and therefore influence rate adaptation. However, there is no evidence that this is an issue as the radio access network links are almost always the bottle neck and not the wireline links.
- It does not provide any end-to-end confirmation that the access networks and UEs all support rate adaptation. This can be addressed as described in clause 5.5.6.1.

### 5.5.5.3 UE Adaptation at Rates > GBR

When a UE risks operating at rates above the GBR it relies on rate adaptation mechanisms to indicate when it has exceeded the supported bandwidth and the UE should reduce its rate. TS 26.114 has multiple mechanisms that can trigger a media receiver to request that the media sender reduce its rate, namely:

- 1) The media receiver experiences packet loss, jitter, or delay in excess of certain internally set thresholds.
- 2) The media receives an indication that the MBR has been reduced below its current transmission rate.
- 3) The media receiver detects packets with ECN-CE markings.
- 4) The media receiver receives an ANBR message indicating that its downlink rate needs to be reduced.

Furthermore, the media sender could also get a direct ANBR from its eNB that it should reduce its uplink transmission rate.

The above mechanisms for rate adaptation are generally arranged in the order of "least graceful" to "most graceful" -- the latter ones could pose the least interruption or degradation to the media.

Therefore, when determining what rate to transmit at above GBR, the UE can adjust the aggressiveness of its algorithms based on knowing which of the above mechanisms are supported by the system, i.e., supported by the access network, the other UE, and the core network.

While ECN support by the far-end UE can be determined via the ECN negotiation procedures, it cannot be confirmed for the near and far-end access networks.

For RAN-assisted rate adaptation, there is currently no apparent way for a UE to confirm that the far-end access network and UE support this form of adaptation. This can be addressed by the solution described in clause 5.5.6.1.

## 5.5.6 Potential Solutions

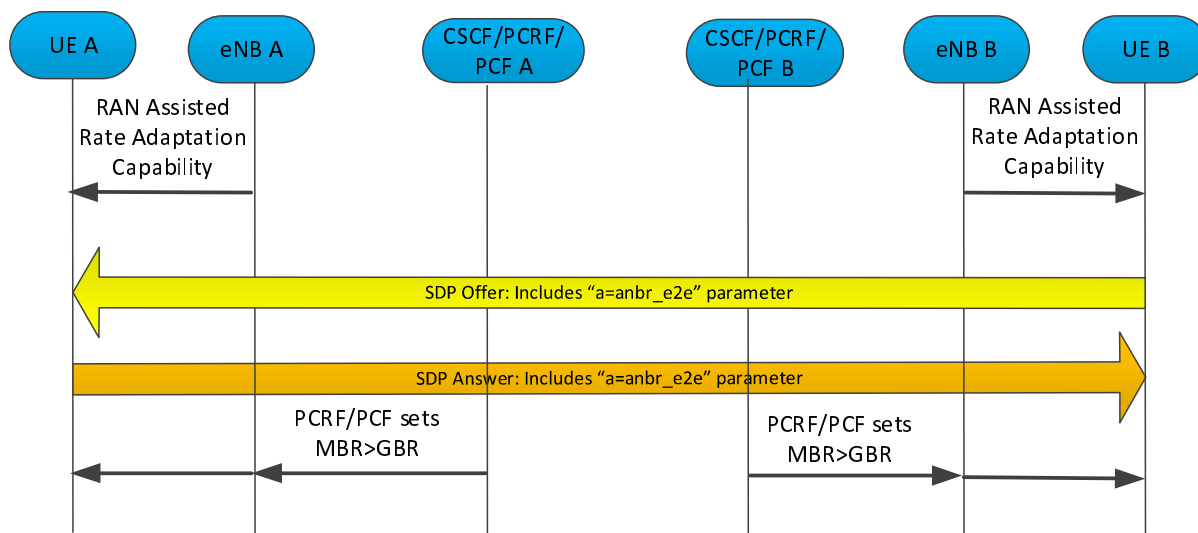
### 5.5.6.1 SDP Parameter for End-to-end RAN-assisted codec adaptation support

Defining a new SDP parameter to indicate that both UEs and their respective eNBs, support RAN assisted codec adaptation would address the issues identified in clauses 5.5.5.2 and 5.5.5.3. The general semantics would be as follows:

- 1) An Offering UE that supports the ANBR messages (RAN-assisted codec adaptation) also checks that its eNB supports the feature. If both support ANBR then the UE includes the *anbr\_e2e* SDP parameter in the SDP offer.
- 2) An Answering UE that receives the *anbr\_e2e* parameter in the SDP Offer includes the *anbr\_e2e* parameter in the SDP Answer if the Answering and its eNB also support the ANBR messages. Otherwise, the Answering UE does not include the *anbr\_e2e* parameter in the SDP Answer.
- 3) An Answering UE that does not receive the *anbr\_e2e* parameter in the SDP offer does not include the *anbr\_e2e* parameter in the SDP answer.

It should be noted that clause 5.1.5.3 of TR 26.910 [13] describes a potential solution based on SDP-based exchange of RAN capabilities using a new 'RANCapabilities' attribute, that not only allows indicating RAN-assisted codec adaptation capabilities, but also other radio capabilities such as delay budget reporting. The newly defined *anbr\_e2e* parameter could be part of the 'RANCapabilities' attribute to indicate RAN-assisted codec adaptation capabilities.

The general steps are illustrated in Figure 5.5.6.1.1.



**Figure 5.5.6.1.1: Negotiation of end-to-end RAN-assisted codec adaptation support.**

The PCRF/PCF would use the presence of this parameter in the SDP Answer to determine that it could set  $MBR > GBR$  with very high confidence and the UE could use its presence to use more aggressive adaptation algorithms when operating at rates  $>GBR$ .

### 5.5.6.2 SDP Parameter for link-by-link RAN-assisted codec adaptation support

A variant on the above is to have the ability for only one UE and its access network to indicate support of the ANBR messages. This may be useful in cases where end-to-end adaptation is not guaranteed but the ability of one link to support it may provide a lower level of confidence to set  $MBR > GBR$  and operate at rates  $> GBR$ .

The semantics of such a solution could be as follows:

- 1) An Offering UE that supports the ANBR message also checks that its eNB supports the feature. If both support ANBR then the UE includes the *anbr\_Off* SDP parameter in the SDP offer.
- 2) An Answering UE that receives the *anbr\_Off* parameter in the SDP Offer includes the *anbr\_Off* parameter in the SDP Answer if the Answering UE and its eNB also support the ANBR messages. Otherwise, the Answering UE does not include the *anbr\_Off* parameter.
- 3) An Answering UE that does not receive the *anbr* parameter in the SDP offer includes the *anbr\_Ans* parameter in the SDP answer only if the Answering UE and its eNB support the ANBR messages.

### 5.5.6.3 Support of RAN-Assisted Codec Adaptation on NR

ECN-based and RAN-assisted codec adaptation are both currently supported for NR. Hence, the PCRF/PCF and UE do not have to purely rely on the rudimentary packet drop, delay, or jitter detection at the UE receiving media to set  $MBR > GBR$  and operate at rates  $> GBR$ , respectively, and can leverage awareness of support for ECN and RAN-assisted codec adaptation when setting  $GBR < MBR$  bearers.

To enable the 5GS to operate at least as well as EUTRAN it is recommended that RAN-assisted codec adaptation also be supported over NR. ECN is not recommended because of the limitations identified in clause 2.

Furthermore, to enable the end-to-end solutions proposed in clauses 4 and/or 5, it is necessary to specify a mechanism for a UE to be able to determine whether its access network supports RAN-assisted codec adaptation prior to sending the first SDP Offer or Answer.

Currently the UE can only determine this when,

- the *bitRateQueryProhibitTimer* is included in a *RRCconnectionReconfiguration* message from the access network; or
- the UE receives a recommended bit rate MAC CE from the access network.

To indicate support of this capability to the UE it is necessary for RAN2 to enhance the RRC specification to explicitly indicate that RAN-assisted codec adaptation is supported by SIB or enabled by RRC Reconfiguration.

## 5.6 Technical Aspect 6: Profiles for 5G Deployments

### 5.6.1 Description

As discussed in clause 4.2 and also in TS 22.261 [2], 5G system is expected to a wide range of applications and market verticals, and media handling in relation to 3GPP conversational services in the context of MTSI can be relevant for these broad set of verticals. Enhanced media handling capabilities are necessary to address certain verticals on the high end (e.g., immersive video/audio for enhanced mobile broadband), and this has resulted in various upgrades to mandatory MTSI client capabilities as documented in the present document, such as those on speech codecs (in clause 5.1), video codecs (in clause 5.2) and media rate adaptation (in clause 5.3). In the meantime, there are other 5G verticals, in which the device complexities need to be kept limited (e.g., wearables, IoT, smartwatch, etc.), or for which superwideband acoustic capability could not be feasible, and media quality expectations could be on par with 4G. For such low-end verticals, the upgrades documented in the present document on mandatory MTSI client media handling features may not be suitable. To address these wide range of media handling capabilities foreseen for emerging 5G usages of MTSI, creation of MTSI client profiles for a selected set of market verticals with a corresponding set of mandatory codecs and other media handling capabilities is desirable.

## 5.6.2 Implications on MTSI

Currently, MTSI specification in TS 26.114 does not include client profiles targeted for various 5G deployments.

## 5.6.3 Implications on IMS-based Telepresence

At this point, no impact is foreseen specifically for IMS-based telepresence in TS 26.223 in regards to this technical aspect.

## 5.6.4 Recommended Requirements

It is recommended that two MTSI client profiles are defined: a default profile that would be the main profile and a constrained profile. This allows some flexibility in TS 26.114 for a selected set of market verticals of 5G deployments. It is also recommended that with each profile, a corresponding set of mandatory codecs and other media handling capabilities are defined. The defined profiles are expected to consider interoperability and target minimizing potential transcoding needs.

## 5.6.5 Gap Analysis

It would be desirable to define MTSI client profiles based on speech and video codec requirements described in clauses 5.2.1.1 and 5.2.2 of TS 26.114. In addition to codecs, other media handling capabilities may be included in the defined MTSI client profiles, but this is not expected to be addressed at the first stage.

## 5.6.6 Potential Solutions

It would be desirable to have two MTSI client profiles in TS 26.114 based on codec capabilities. One of these profiles would be based on Rel-15 client capabilities described in clauses 5.2.1.1 and 5.2.2 of TS 26.114. The second MTSI client profile would be based on speech and video codec capabilities relevant to serve the needs of low-end 5G verticals such as wearables and IoT. A candidate profile for this purpose could adopt Rel-14 client capabilities described in clauses 5.2.1.1 and 5.2.2 of TS 26.114.

# 5.7 Technical Aspect 7: Support for Real-Time Interaction

## 5.7.1 Description

Real-time communication today consists of voice, video and text. We utilize these communication services to write to each other, talk to each other and see each other in real-time. With the introduction of 5G, it should be considered how we can take the communication service to the next level.

The industry could take the next step and evolve this infrastructure and services further for a 5G world, adding on new capabilities building on IMS.

With 4G we have real-time communication, the ability to see, hear and talk over the network. With 5G and the resulting bandwidth and low latency we could add other abilities such as touch and move over the network, extending real-time communication with real-time interaction.

The IMS platform inherently provides global find-and-connect with phone number, authentication, quality-of-service, mobility, security, and robustness; all of them qualities that are essential for real-time critical point-to-point communications. The low latency of 5G will enable new types of use cases, where real-time communication and interaction will be possible over long distances. This is a significant change and will potentially facilitate a global shift from real-time communication to real-time interaction. Technology-wise this can be enabled by adding a data channel to IMS, alongside with voice, video and text channels.

This shift may create new revenue opportunities for the communication service providers in various industry segments and in some consumer areas as well. This potentially unlocks an addressable market for communication service providers worldwide in areas like: Remote Health, Remote Drone Control, Remote Education, Remote Security, Remote Machinery Control, Augmented Reality, and Consumer-to-business Interaction.

As an example of a remote health use case that could be enabled by such data channel, today, ambulances and paramedics have limited capability to engage remote healthcare professionals and therefore must rely on the expertise of the paramedic. Few have the training or equipment to perform ultrasound diagnosis. If the doctor can be available remotely for the patient, the right diagnosis can be performed upfront with the right decisions on which hospital the patient should be transferred to, and what preparation is needed ahead of arrival. This can save precious time and money for the Health Care Provider.

This use case involves that a patient is investigated with ultrasound by the paramedic in the ambulance, with remote support from a doctor in a hospital far away. Utilizing standard IMS for securing point to point HD voice, video calling and data communication (over an IMS data channel) a paramedic that requires remote healthcare assistance can connect a voice and video call for 2-way communication. When the paramedic determines the patient would benefit from an ultrasound diagnosis, the paramedic can place a video call to the centralized ultrasound unit. There will be a live video link between the doctor and the patient, allowing the doctor to see the patient and give some commands, like "turn left", "breathe in", etc.

With the use of an IMS data channel, they can now also extend the ultrasound equipment in the ambulance to the doctor for remote control, with the doctor being able to see and manage the ultrasound machine directly. The doctor conducts a remote ultrasound procedure in real-time by directing the paramedic that is wearing a glove with haptic feedback connected through the IMS data channel. The connected haptic glove is controlled by the doctor through a joystick control sending small vibrations to the paramedic's glove to direct the ultrasound sensor, whilst the doctor studies the image from the ultrasound equipment that is sent on another IMS data channel. This allows for a nonverbal, tactile interaction between doctor, paramedic and patient. The paramedic becomes the extension of the doctor in interacting directly with the patient, and the doctor is guiding the paramedic through the diagnosis. The doctor, paramedic and patient can also discuss and interact to determine next steps. Following diagnosis, the choice of hospital in which to send the patient saves valuable time and ambulance resources, and allows the team receiving the patient to prepare ahead of arrival.

### 5.7.2 Implications on MTSI

Currently, MTSI specification in TS 26.114 does not include any possibility for such data channel usage.

### 5.7.3 Implications on IMS-based Telepresence

Currently, IMS-based telepresence in TS 26.223 has a suitable data channel, but that is only used to exchange CLUE configuration and messages and support would have to be generalized similar to MTSI (above).

### 5.7.4 Recommended Requirements

To enable the above use case and other real-time interaction use cases in the IMS communication service framework, it is recommended that the following additions are made to IMS, MTSI, and IMS Telepresence (not in any priority order):

- 1) Secure, in-call, UE-to-UE transfer of application-specific information, to meet as many end-to-end, real-time interaction needs as possible
- 2) No need to standardize or specify every application using the IMS data channel, for maximum usefulness, fast innovation, and fast deployment
- 3) Easy access to data channel input and output for the application and end-user, for fast application development and maximum usefulness
- 4) Multiple, IMS data channel applications can be pre-defined and used per UE
- 5) Which IMS data channel application(s) to use can be chosen on a per-call basis
- 6) Each IMS data channel application can use multiple data channels simultaneously within a single call
- 7) Avoid the need to create an entirely new ecosystem for IMS data channel application development, to limit the development and deployment effort for both operators and developers
- 8) High flexibility in IMS data channel transport characteristics, to support as many applications and usages as possible

- 9) Allow for operator control of applications and application use of resources, to prevent misuse of IMS and operator assets

### 5.7.5 Gap Analysis

It would be desirable to define MTSI data channel as another media in TS 26.114 and generalize data channel usage in TS 26.223.

### 5.7.6 Potential Solutions

#### 5.7.6.1 WebRTC Data Channel Framework

##### 5.7.6.1.1 Flexibility in Usage

Through the capabilities delivered with IMS, use of the WebRTC data channel framework [23][24] would allow for configurable IMS data channel definitions whilst still ensuring quality-of-service, security and robustness. These data channels can be used to transport any types of data for any purpose, all in sync with the voice and video communication session. The data channels could be established with operator provided characteristics for the data to be sent based on existing IMS policy framework, to meet the requirements of specific use cases when it comes to latency, robustness and bandwidth. The WebRTC data channel setup can be negotiated in SDP [25] [26], which is a key component of IMS, and WebRTC data channels can therefore be an almost seamless addition to an existing IMS voice/video call. The choice of WebRTC data channel as technology in the context of IMS would therefore tentatively meet requirements 1, 6, and 8 above.

##### 5.7.6.1.2 Flexibility in Development

Requirements 2, 3, and 7 can be harder to meet in a traditional IMS context and to meet them requires careful consideration, especially on the UE side.

Starting from requirement 7, if IMS data channels are based on WebRTC data channels, there's already substantial WebRTC data channel support in all major web browsers, which also means that all major web development tools today include means to make use of WebRTC data channel. This involves use of the JavaScript programming language that can be tied to the HTML code of a web page, which is a common technique used today for any non-static web content. Available development tools include, for example, the developer mode of the web browsers themselves, as well as modern web server-side tools such as "node.js" [27], which would likely mean that existing web developers would find application development for IMS data channel very familiar.

As for requirements 2 and 3, use of JavaScript and (HTML) web pages provides a very good starting point; new applications using the IMS data channel can easily be created in JavaScript without the need to change anything around the IMS data channel itself. Having the JavaScript interpreter tied to the device platform web engine additionally provides easy programmatic access to most (if not all) UE input/output devices that JavaScript applications can use to produce and consume data sent through the IMS data channel. In terms of the above listed use case as an example, the ultrasound image from the ultrasound machine, the screen showing that ultrasound image to the doctor, the joystick, and the haptic glove are input/output devices attached to IMS data channels through physical interfaces on the UEs and must be handled by the IMS data channel JavaScript application.

Another benefit that comes from using the UE platform web engine is that programmatic JavaScript access to textual and graphical user interface elements is already solved through the existing web engine JavaScript control over any imaginable web page element. Since one of the main potential strengths of IMS data channel is that it can be used for almost anything, it wouldn't work to require a hard-coded user interface that must be used by all IMS data channel applications. The IMS data channel JavaScript application needs flexible user interface input and output to handle end-user interaction and presentation related to IMS data channel data. This would however also suggest that the legacy dialer, traditionally used for IMS communication services, must be amended with capability to handle interactive web content in the form of HTML and JavaScript, to keep IMS data channel handling easily reachable within an IMS communication context.



### 5.7.6.1.3 Flexibility in Operations

A few remaining solution aspects on requirement 2 come from the desire to easily handle multiple IMS data channel applications in requirements 4, 5, and introducing IMS data channel application operations and control in requirement 9:

- It can be seen from the above that it is beneficial to be able to use any available web development tool, and it is therefore assumed that the web page and JavaScript used for IMS data channel are created in that way. It can further be assumed that the IMS data channel input/output and the corresponding graphical user interface (being integrated with the native dialer) must be more restricted than the general web context, which would require some verification and potential modification of the web page and the JavaScript produced by the general web development tools, before deployment in IMS data channel context.
- The WebRTC use of data channel assumes that the two end-points have access to the same JavaScript, which contains the specific application logic on how this data channel is to be used and its interaction with the end-user through the graphical user interface. It is suggested to not step away from this principle, which means that the IMS data channel web page and JavaScript must be available in both UE before use.
- It is assumed that a single end-user may want to use the same UE with several different IMS data channel applications, perhaps depending on what the desired IMS data channel use case is, perhaps also depending on who the remote party is, which means that IMS data channel application must at least be possible to change on a per-call basis. It is therefore assumed that a single UE must have access to some repository of IMS data channel applications that the end-user can select from. Such application repository could be stored locally in the UE, somewhere in the public cloud, or as an operator-provided repository that is likely linked to the IMS data channel subscription.

Three benefits with operator-provided IMS data channel application repository can be noted:

- a) Ability to combine the upload process with adapting the provided JavaScript and web page according to a format that will suit the UE dialer graphical layout and IMS data channel usage, as opposed to how they can be used for legacy browser WebRTC data channel on the Internet.
- b) Ability to dynamically distribute an IMS data channel application that one UE desires to use in the call to the other party's UE, as part of the call itself, simplifying usage and avoiding separate, pre-call procedures that might be considered cumbersome by the end-users and hamper IMS data channel usage.
- c) Increased possibilities for the end-user to easily choose own IMS data channel application on a per-call basis, during or just before the IMS call setup (remote user might get a data channel application pushed to it as part of b) above).

### 5.7.6.1.4 Relation to Existing WebRTC in IMS Specifications

There is a specification of WebRTC access to IMS in 3GPP [31]. That specification is however focused on access to IMS from an IMS-adapted WebRTC client or using IMS as a communication network between IMS-adapted WebRTC clients, effectively extending IMS to the WebRTC domain. Clause 8 of that specification describes use of the WebRTC data channel, but only to pass data between WebRTC clients adapted to IMS usage. Additionally, IMS will only pass through the content of such WebRTC data channel if the eIMS-AGW can translate between WebRTC data channel and some other protocol that is supported inside the IMS domain, e.g. MSRP where MSRP is carried in the WebRTC data channel [30] outside of IMS domain. Therefore, extending IMS itself to handle WebRTC data channel as another media alongside voice, video, and text seems to be out of scope for that specification today.

### 5.7.6.2 MSRP

It could be argued that other data channel capabilities are already available in IMS, such as e.g. MSRP [28][29]. However, MSRP is mainly a messaging protocol and is not designed for stream data. Also, MSRP doesn't have the tight relation to HTML and JavaScript provided by WebRTC data channel and the benefits from that is typically missing from MSRP implementations, which make them much less flexible and less easy to use for data channel application developers. While it would be possible to expose a platform IMS MSRP stack to an application on the UE in the same fashion as the WebRTC data channel is exposed to JavaScript, no such plans are currently known. It should be noted that choosing WebRTC data channel as technology for IMS data channel doesn't prevent use of MSRP; In case MSRP is needed by future IMS data channel applications, e.g. to interwork with existing MSRP applications, work is ongoing in IETF to describe how to run MSRP within the WebRTC data channel [30].

---

## 6 Impacts of 5G Stage-2 System Architecture on 3GPP Conversational Services

### 6.1 Technical Aspect 1: Mapping of Conversational Services to 5G System

#### 6.1.1 Description

Clause 4.3 describes the Stage-2 architecture for the 5G system. The 5G system is to support 3GPP conversational services including MTSI and IMS-based telepresence and it is of interest to understand the mapping of these services to the 5G system. In that regard, the control-plane and user-plane functionality of a network slice on MTSI and IMS-based telepresence are of interest and how these would map to the 5G system architecture.

#### 6.1.2 Implications on MTSI and IMS-based Telepresence

The use of 5G System to access IMS in Rel-15 has been specified in TS 23.228 Annex Y [7]. In addition, PCC aspects are described in TS 23.503 [8].

As described in the specifications referenced above, for Rel-15 it is assumed that the IMS can be accessed via 5G System without any changes to the IMS. The IMS subsystem connects to the underlying 5G System using the Sh/Cx interface towards the UDM/HSS and the N5 or Rx interface to the PCF.

The N5 reference point is a point-to-point representation of the service-based (HTTP-based) interface Npcf/Naf. While both N5 and Rx can be used for communication between the IMS and the PCF, it is noted that in Rel-15 the two reference points are functionally equivalent.

IMS functional blocks relevant for MTSI and IMS-based Telepresence services, including Application Server (AS), S-CSCF, I-CSCF, P-CSCF, HSS, MRFC, MRFP, IBCF, TrGW, etc. are considered as part of the AF and are transparent to the 5G System as long as the reference points N5/Rx, Sh and Cx remain unchanged. Furthermore, the 5G system architectural requirements specified in clause 5.16.3 of TS 23.501 [3] on IMS support impacting AMF, SMF and PCF are applicable for MTSI.

---

## 7 Codecs for 5G Conversational Services

### 7.1 Introduction

Clause 5 of the present document describes some codec aspects for 5G conversational services. As described there, reuse of codecs that are already part of 3GPP conversational services seems sufficient to meet most 5G requirements. This clause describes new codecs and new usage of existing codecs that are enabled by the 5G system.

---

## 8 QoS Handling for 5G Conversational Services

### 8.1 Introduction

No gaps were identified in regards to QoS handling for MTSI and IMS-based telepresence over the 5G system.

---

## 9 Conclusions

It is recommended to conduct normative work toward specifying the following functionality in TS 26.114 and TS 26.223:

- For video, mandate support for H.265/HEVC and H.264/AVC for 5G MTSI and IMS Telepresence endpoints, as per the recommendations in clause 5.2.4.
- For speech, mandate support for AMR, AMR-WB, and develop further codec requirements based upon EVS, potentially including SWB operation, as per the recommendations in clause 5.1.4.
- For media rate adaptation, recommend support for speech and video adaptation capabilities considering the potential solutions described in clause 5.3.5.
- For 5G NR access, enable support for media handling with NR access, including that for speech and video, toward addressing the gaps indicated in clause 5.5.5. This also includes support for negotiation of radio capabilities on ANBR signalling as described in clause 5.5.6.
- For profiles for 5G deployments: define MTSI client profiles with a corresponding set of mandatory codec and potentially other media handling capabilities to address the needs and constraints of different terminal categories (e.g. IoT, wearables) related to different 5G verticals, considering the potential solutions described in clause 5.6.6 of TR 26.919. Profiles will be defined in a clear way in relation to identified terminal characteristics.

---

## Annex B:

### Change history

Change history							
Date	Meeting	TDoc	CR	Rev	Cat	Subject/Comment	New version
06-2018	80	SP-180282				Presented to TSG SA#80 (for information)	1.0.0
09-2018	81	SP-180959				Presented to TSG SA#81 (for approval)	2.0.0
09-2018	81					Approved at TSG SA#81	16.0.0
03-2019	83	SP-190039	0001	-	F	ECN Support in NR	16.1.0
09-2019	85	SP-190650	0002	-	B	Addition of 5G Real-Time Interaction	16.2.0

---

# History

Document history		
V16.2.0	November 2020	Publication