3GPP TR 22.823 V16.1.0 (2018-06)

Technical Report

3rd Generation Partnership Project;

Technical Specification Group Services and System Aspects;

Study on Enhancements to IMS for   
new real time communication Services;

Stage 1

(Release 16)

** 

The present document has been developed within the 3rd Generation Partnership Project (3GPP TM) and may be further elaborated for the purposes of 3GPP.  
The present document has not been subject to any approval process by the 3GPPOrganizational Partners and shall not be implemented.  
This Report is provided for future development work within 3GPPonly. The Organizational Partners accept no liability for any use of this Specification.  
Specifications and Reports for implementation of the 3GPP TM system should be obtained via the 3GPP Organizational Partners' Publications Offices.

Keywords

IMS, real time

***3GPP***

Postal address

3GPP support office address

650 Route des Lucioles - Sophia Antipolis

Valbonne - FRANCE

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

Internet

http://www.3gpp.org

***Copyright Notification***

No part may be reproduced except as authorized by written permission.  
The copyright and the foregoing restriction extend to reproduction in all media.

© 2018, 3GPP Organizational Partners (ARIB, ATIS, CCSA, ETSI, TSDSI, TTA, TTC).

All rights reserved.

UMTS™ is a Trade Mark of ETSI registered for the benefit of its members

3GPP™ is a Trade Mark of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners  
LTE™ is a Trade Mark of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners

GSM® and the GSM logo are registered and owned by the GSM Association

Contents

Foreword [5](#__RefHeading___Toc517529769)

1 Scope [6](#__RefHeading___Toc517529770)

2 References [6](#__RefHeading___Toc517529771)

3 Definitions, and abbreviations [6](#__RefHeading___Toc517529772)

3.1 Definitions [6](#__RefHeading___Toc517529773)

3.2 Abbreviations [7](#__RefHeading___Toc517529774)

4 Use Cases [7](#__RefHeading___Toc517529775)

4.1 One source to many destinations communication [7](#__RefHeading___Toc517529776)

4.1.1 Description [7](#__RefHeading___Toc517529777)

4.1.2 Pre-conditions [7](#__RefHeading___Toc517529778)

4.1.3 Service Flows [7](#__RefHeading___Toc517529779)

4.1.4 Post-conditions [8](#__RefHeading___Toc517529780)

4.1.5 Potential Impacts or Interactions with Existing Services/Features [8](#__RefHeading___Toc517529781)

4.1.6 [Potential] Requirements [8](#__RefHeading___Toc517529782)

4.2 Remote camera control [8](#__RefHeading___Toc517529783)

4.2.1 Description [8](#__RefHeading___Toc517529784)

4.2.2 Pre-conditions [8](#__RefHeading___Toc517529785)

4.2.3 Service Flows [8](#__RefHeading___Toc517529786)

4.2.4 Post-conditions [9](#__RefHeading___Toc517529787)

4.2.5 Potential Impacts or Interactions with Existing Services/Features [9](#__RefHeading___Toc517529788)

4.2.6 [Potential] Requirements [9](#__RefHeading___Toc517529789)

4.3 Fast call set up time for Group Communication [9](#__RefHeading___Toc517529790)

4.3.1 Description [9](#__RefHeading___Toc517529791)

4.3.2 Pre-conditions [9](#__RefHeading___Toc517529792)

4.3.3 Service Flows [9](#__RefHeading___Toc517529793)

4.3.4 Post-conditions [9](#__RefHeading___Toc517529794)

4.3.5 Potential Impacts or Interactions with Existing Services/Features [9](#__RefHeading___Toc517529795)

4.3.6 Potential Requirements [10](#__RefHeading___Toc517529796)

4.4 Augmented reality [10](#__RefHeading___Toc517529797)

4.4.1 Description [10](#__RefHeading___Toc517529798)

4.4.2 Pre-conditions [10](#__RefHeading___Toc517529799)

4.4.3 Service Flows [10](#__RefHeading___Toc517529800)

4.4.4 Post-conditions [10](#__RefHeading___Toc517529801)

4.4.5 Potential Impacts or Interactions with Existing Services/Features [10](#__RefHeading___Toc517529802)

4.4.6 [Potential] Requirements [10](#__RefHeading___Toc517529803)

4.5 VR Telepresence [11](#__RefHeading___Toc517529804)

4.5.1 Description [11](#__RefHeading___Toc517529805)

4.5.2 Pre-conditions [11](#__RefHeading___Toc517529806)

4.5.3 Service Flows [11](#__RefHeading___Toc517529807)

4.5.4 Post-conditions [12](#__RefHeading___Toc517529808)

4.5.5 Potential Impacts or Interactions with Existing Services/Features [12](#__RefHeading___Toc517529809)

4.5.6 [Potential] Requirements [12](#__RefHeading___Toc517529810)

4.6 IMS network slicing [13](#__RefHeading___Toc517529811)

4.6.1 Description [13](#__RefHeading___Toc517529812)

4.6.2 Pre-conditions [14](#__RefHeading___Toc517529813)

4.6.3 Service Flows [14](#__RefHeading___Toc517529814)

4.6.4 Post-conditions [14](#__RefHeading___Toc517529815)

4.6.5 Potential Impacts or Interactions with Existing Services/Features [14](#__RefHeading___Toc517529816)

4.6.6 [Potential] Requirements [14](#__RefHeading___Toc517529817)

4.7 Service visibility and usage [15](#__RefHeading___Toc517529818)

4.7.1 Description [15](#__RefHeading___Toc517529819)

4.7.2 Pre-conditions [15](#__RefHeading___Toc517529820)

4.7.3 Service Flows [15](#__RefHeading___Toc517529821)

4.7.4 Post-conditions [15](#__RefHeading___Toc517529822)

4.7.5 Potential Impacts or Interactions with Existing Services/Features [15](#__RefHeading___Toc517529823)

4.7.6 [Potential] Requirements [15](#__RefHeading___Toc517529824)

4.8 Services using Artificial Intelligent (voiceprint) in real time communication [15](#__RefHeading___Toc517529825)

4.8.1 Description [15](#__RefHeading___Toc517529826)

4.8.2 Pre-conditions [16](#__RefHeading___Toc517529827)

4.8.3 Service Flows [16](#__RefHeading___Toc517529828)

4.8.4 Post-conditions [16](#__RefHeading___Toc517529829)

4.8.5 Potential Impacts or Interactions with Existing Services/Features [16](#__RefHeading___Toc517529830)

4.8.6 [Potential] Requirements [16](#__RefHeading___Toc517529831)

4.9 Services using Artificial Intelligent ( on demand translation) in real time communication [16](#__RefHeading___Toc517529832)

4.9.1 Description [16](#__RefHeading___Toc517529833)

4.9.2 Pre-conditions [17](#__RefHeading___Toc517529834)

4.9.3 Service Flows [17](#__RefHeading___Toc517529835)

4.9.4 Post-conditions [17](#__RefHeading___Toc517529836)

4.9.5 Potential Impacts or Interactions with Existing Services/Features [17](#__RefHeading___Toc517529837)

4.9.6 [Potential] Requirements [17](#__RefHeading___Toc517529838)

5 Considerations [17](#__RefHeading___Toc517529839)

5.1 Considerations on security [17](#__RefHeading___Toc517529840)

5.2 Considerations on Charing Aspect [17](#__RefHeading___Toc517529841)

6 Conclusion and Recommendations [18](#__RefHeading___Toc517529842)

Annex A: Change history [19](#__RefHeading___Toc517529843)

# Foreword

This Technical Report has been produced by the 3rd 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.

# 1 Scope

The present document aims to develop high-level use cases and identify the related potential requirements to enable IMS to support new real time communication services.

IMS enhancements for the following scenarios are studied in this document:

* one to many communication (e.g. live broadcast video service in a stadium, concert, etc.) with efficient media negotiation;
* interworking for a UE to communicate with non-sip devices (e.g. camera), and delivery of information to control functions of the non-sip devices (e.g. PTZ);
* fast IMS call setup time for Group Communication;
* enhancements on voice & video communications for supporting AR/VR;
* network slicing for IMS;
* service visibility and usage;.
* services using AI (voiceprint) in RTC.

# 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 the 5G system: Stage 1”.

[3] 3GPP TS 22.468: “Group Communication System Enablers for LTE”.

[4] 3GPP TS 22.243: "Speech recognition framework for automated voice services".

[5] 3GPP TR 22.977: "Feasibility study for speech-enabled services".

# 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].

**<defined term>:** <definition>.

… …

## 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].

AR Augmented Reality

ASR Automatic Speech Recognition

ONVIF Open Network Video Interface Forum

PTZ Pan, Tilt and Zoom

SRF Speech Recognition Framework

VoNR Voice over New Radio

# 4 Use Cases

## 4.1 One source to many destinations communication

### 4.1.1 Description

Fully mobile and connected society will need efficient distribution of information from one source to many destinations. These services may distribute content as done today (typically only downlink), but also provide a feedback channel (uplink) for interactive services or acknowledgement information. Both, real-time or non-real time services should be possible. Furthermore, such services are well suited to accommodate vertical industries’ needs. These services are characterized by having a wide distribution which can be either geo-location focused or address-space focused (many end-users).

Beyond 2020, receiving text/pictures, audio and video, everywhere and as soon as things happen (e.g., action or score in a football match) will be common. Stadium services, advertisements, voucher delivery, festivals, fairs, and congress/convention are typical scenarios.

For one source to many destinations voice/video communication, typical scenario is that most of the audiences may only have downlink media while a small portion of audiences may have bidirectional real time interaction. A mechanism is needed to manage the priority/rights of users, e.g. interactive or receive-only, performance (e.g. latency) and to deliver the content efficiently to the users. In-network content caching provided by the operator, a 3rd party or both, can improve user experience, & reduce bandwidth pressure.

### 4.1.2 Pre-conditions

Alice, Bob, Charlie, David are customers who can simultaneously receive video stream from live broadcast event. For this broadcast Bob is the chairman. Alice is authorised to have interactive service, while Charlie and David can only receive the content.

### 4.1.3 Service Flows

Bob invite Alice, Charlie and David to join the live broadcast (indicates how to join the live broadcast and the address e.g. URL). Alice is authorized to have interactive service, while Charlie and David only receive the content. The role/mode of participants can be changed after authorization and is managed by the IMS network (e.g., receive-only, interactive).

After receiving the content from the broadcast server, the network delivers it to caching entities which are close to the users.

Charlie requests to change and join the live broadcast in interactive mode, which is authorized by Bob.

During the broadcast, Bob authorizes Alice to take over the chairmanship.

David requests to change and join the live broadcast in interactive mode, which is authorized by Alice.

Chairman terminates the live broadcast.

### 4.1.4 Post-conditions

Users are able to watch the video from a content caching entity close to them, and have interactive communication according to their priority/rights.

### 4.1.5 Potential Impacts or Interactions with Existing Services/Features

None identified.

### 4.1.6 [Potential] Requirements

In addition of the requirements in TS 22.261[2] Clause 6.6 Efficient content delivery, the IMS network shall support the following requirements.

For one-to-many communication service, the IMS network shall provide a mechanism to support a single session with large number of users (e.g. more than 10K) that can be in different communication modes e.g. interactive or receive only.

For one-to-many communication service, the IMS network shall support the different communication modes (e.g., receive-only, interactive, or switch among these modes).

The IMS network shall be able to control the communication mode of a UE based on the indication from a UE that acts as the chairman of the session (i.e. based on UE role).

The IMS network shall authorize a user to join a single session served for one-to-many communication service.

The IMS network shall be capable of deleting a user from the session served for one-to-many communication service.

For one-to-many communication service, the IMS network shall provide a mechanism to allow an application to allocate efficiently manage resources (e.g. backhaul resources and/or application resource) for large numbers of users in a single session to provide good quality of experience.

For one-to-many communication service, the IMS network shall be able to support designated coverage area. Only users within the coverage area receive the one-to-many communication service.

## 4.2 Remote camera control

### 4.2.1 Description

Currently, video client (e.g. camera) has been widely deployed in various vertical industries. Real time communication between UE and camera is emerging, e.g. for ambulances equipped with high-definition cameras, doctors in the hospital can give remote guidance about medical treatment / diagnosis through mobile phone and control the camera in real-time for better footage.

For this scenario the key aspect is that IMS provides service (interworking) to a UE that allow one protocol (e.g. SIP) to control non-sip devices (e.g. camera) that support different protocols (e.g. ONVIF).

### 4.2.2 Pre-conditions

Camera in the ambulance connects to the operator’s network.

Doctor’s UE connects to the operator’s network.

### 4.2.3 Service Flows

Doctor requests communication with the camera to provide real-time medical treatment.

Doctor requests control of the camera (e.g. Pan, Tilt, and Zoom (PTZ)) to gain better footage.

### 4.2.4 Post-conditions

Effective medical treatment is given.

### 4.2.5 Potential Impacts or Interactions with Existing Services/Features

None identified.

### 4.2.6 [Potential] Requirements

The IMS network shall provide a mechanism to allow a user to discover a video client (e.g. camera).

The IMS network shall support the interworking (protocol conversion) for a UE to communicate with a non-sip device (e.g. camera).

The IMS shall be able to deliver the information to control the functions of a non-sip device (e.g. PTZ of camera).

NOTE: Existing protocols (e.g. ONVIF) could be used.

## 4.3 Fast call set up time for Group Communication

### 4.3.1 Description

As stated in TS 22.468 [3]:

*The system should provide a mechanism to support a Group Communication end-to-end setup time less than or equal to 300ms. It is assumed that this value is for an uncontended network, where there is no presence checking and no acknowledgements requested from Receiver Group Member(s). The end-to-end setup time is defined as the time between when a Group Member initiates a Group Communication request on a UE and the point when this Group Member can start sending start sending a voice or data communication.*

*The time from when a UE requests to join an ongoing Group Communication to the time that it receives the Group Communication should be less than or equal to 300ms.*

*Note: The 300 ms indicated in the preceding requirements is based on requirements from ETSI ETR 086 [8] for legacy TETRA mission critical voice systems. It is understood that* *these requirements are particularly important for half duplex voice communication and other data that is delay sensitive. These requirements may not be met in some cases where the data is delay insensitive e.g., a large document and/or where the type of Group Communication requires acknowledgement(s) from Receiver Group Members before it is allowed to proceed.*

Although Group Communication is supported in IMS, the above requirement for the end-to-end setup time of less than or equal to 300ms is not fulfilled. IMS network needs to be enhanced to support critical group communication, like MCPTT.

### 4.3.2 Pre-conditions

None.

### 4.3.3 Service Flows

None.

### 4.3.4 Post-conditions

None.

### 4.3.5 Potential Impacts or Interactions with Existing Services/Features

None identified.

### 4.3.6 Potential Requirements

The IMS shall be able to support a Group Communication end-to-end setup time of less than or equal to 300ms.

NOTE: The end-to-end setup time is defined as the time between when a Group Member initiates a Group Communication request on a UE and the point when this Group Member can start sending a voice or data communication.

NOTE: Roaming case is out of scope.

## 4.4 Augmented reality

### 4.4.1 Description

Currently, Conference (CONF) is one of the IMS Multimedia Telephony supplementary services. As new technologies are emerging, more and more operators hope to enrich the user experience based on Augmented Reality (AR) technology, virtual meeting allows users to share and enhance users’ experiences in real-time with the same quality of communication as in a face-to-face meeting. IMS can be used to set up and manage the AR media streams to ensure a good user experience, i.e. users in the meeting enjoy a strong sense of realism and presence between all participants. Low latency and high bandwidth of communication are needed for AR service.

### 4.4.2 Pre-conditions

The IMS network supports Augmented Reality (AR) service.

Operator provides the AR service (e.g. AR meeting) for the users.

### 4.4.3 Service Flows

Bob books an AR meeting.

Bob, Alice, Mary and David join the AR meeting. Bob and Alice are in the same meeting room, while Mary and David are in another meeting room in different location.

The server of IMS network sets up the connection among Bob’s, Alice’s, Mary’s and David’s terminals, and the AR meeting data are transferred among all the participants via the operator’s network.

Participants’ position will be tracked and affect the synthesis of sounds, thus, every participant can hear positional sound (i.e. character of voice plus positions). .

Each participant’s head motion will be tracked and affect the user visible stereo video.

### 4.4.4 Post-conditions

Bob, Alice, Mary and David who attend the video meeting enjoy a good sense of realism and presence between all participants.

### 4.4.5 Potential Impacts or Interactions with Existing Services/Features

None identified.

### 4.4.6 [Potential] Requirements

The IMS network shall be able to provide the required QoS (e.g., reliability, latency, and bandwidth) for AR service.

To support AR service, the IMS network shall be able to support the motion-to-photon latency requirement as specified in TS 22. 261, section 7.2.3.

The IMS network shall be capable of monitoring the QoS provided for AR media stream for assurance of the AR service quality.

The IMS network shall be able to inform the AR application layer when the required QoS cannot be provided.

The IMS network should be able to terminate different media streams individually (e.g. keep only the sound stream of AR media when the required QoS for AR media stream cannot be provided).

The IMS network shall be able to support speech coding for positional sound.

The IMS network shall be able to set up and manage AR media stream (e.g. coding and de-coding of video stream, composition of virtual and real world objects).

Editor’s Note: Operations over multi-operator networks and large distances may have impact on the AR service.

## 4.5 VR Telepresence

### 4.5.1 Description

The general concept of a VR Telepresence system is to broadcast your own environment (e.g. as 360 degree video) and allow another party to enter your environment, while being able to communicate and interact with you. The following use case shows such a communication link.

Anne is currently on a business in Japan and wants to join a meeting with her team back home in Amsterdam. The team is located in a meeting room around a 360 camera and a microphone array is located on top of the camera to capture spatial sound. The conferencing application on Anne’s tablet shows a section of the 360-degree view of the conferencing room. On voice activity, the camera shows the person currently talking. Anne can swipe the image to look into any direction in the conference room. Equally Anne could switch to an VR head mounted display (HMD) or an AR glasses to experience a better immersion and presence of the meeting room in Amsterdam.

In this scenario one side captures a 360-degree spherical video (e.g. accomplished through the use of an omnidirectional camera) and 3D spatial audio (e.g. captured with a microphone arrays). This captured image can be shown on a traditional display or within an HMD (for full immersion). However, if Anne’s mobile device is not supporting the capability to encode the full 360-degree video, or if Anne has limited bandwidth (e.g. due to being in a crowded office area), the IMS system needs to only send Anne’s end device a part of the video (i.e. the current viewpoint of Anne). Such re-encoding needs to be executed on a server physically as close as possible to Anne’s current location (i.e. in Japan). Otherwise the motion-to-photon delay would be too high to adopt the view quick enough and Anne’s experience would suffer.

Extending on the use case from above, another colleague of Anne, Paul, is joining the conversation from Singapore. In this case the IMS system has to manage the streams between multiple user devices. Thus, the IMS system might implement intelligent routing strategies like Audio/Video MCUs (multipoint control units), in order to reduce the bandwidth load on the overall system. As well as the re-encoding component such MCUs also depend on the physical location of the end users.

### 4.5.2 Pre-conditions

The conference room of the team in Amsterdam is equipped with a 360-degree camera and a microphone array that allow the recoding of spatial audio.

Both Anne and Paul need to have a 3GPP VR capable end device displaying the 360-degree video either in a HMD or on a traditional screen. In case of a traditional screen the device needs sufficient UI capabilities to allow a user to navigate the 360-degree video.

Further the devices of Anne and Paul need to be able to render spatial audio based on the current viewport of the user in the 360 degree space.

### 4.5.3 Service Flows

Alice and the Team are joining the immersive call.

Negotiating the audio / video stream transmission, the IMS system identifies a re-encoding component physically close to Anne.

The video transmission from the team is send into the network as full 360-degree image re-encoded and send as Anne’s current viewport to Anne’s end device.

Further a synchronized special audio stream is send from the Amsterdam conference Room to Anne.

The transmission from Alice is a traditional audio/video link.

### 4.5.4 Post-conditions

Alice enjoys an immersive experience, while the team can freely communicate (including Alice being part of the conversation).

### 4.5.5 Potential Impacts or Interactions with Existing Services/Features

None identified.

### 4.5.6 [Potential] Requirements

The IMS system shall support sufficient SDP-based mechanisms for the negotiation of streaming and VR capabilities across senders and receivers during both call setup and mid-call

The IMS system shall support spatial audio, in this case it has to be transmitted in sync and via multiple channels (to support 3D and spatial characteristics)

The IMS system shall support metadata (e.g. location of users) to orchestrate the communication links and media servers (e.g. a viewport generators, re-encoders) involved in the session.

The IMS system shall support different audio modes, e.g.:

- No audio (audio is not transmitted or via a different channel, e.g. to support lower latency for audio vs. video);

- Stereo audio;

- Multi-channel audio in sync with the video and with spatial information.

The IMS system shall support different delay modes, e.g.:

- Ultra-Low latency mode, with a camera to display delay that is as close to that of MTSI as possible (to enable conversational services);

- Moderate latency mode, with delay of <10sec (for one way immersive media content, like remote teaching);

- High latency mode, with a delay of 10-30sec (for video broadcasting media, like stadium view, concert, webinar).

The IMS system shall support different (immersive) video quality modes, e.g.:

- Low quality mode, e.g. 2k video resolution with 30fps (up to 15Mbps);

- Moderate quality mode, e.g. 4k video resolution with 60fps (up to 50Mbps);

- High quality mode, e.g. 8k+ video resolution with 90+ fps (up to 339Mbps).

The IMS system shall signal the different Audio, Video and Delay modes.

## 4.6 IMS network slicing

### 4.6.1 Description

IMS Network slicing allows the operator to provide customised IMS networks, based on e.g. service categories, service performance requirements.

An IMS network slice can be composed of all or part of IMS network functional entities.

One IMS network can support one or several network slices.

IMS network slice along with 5GC slice provide an end to end customised network for IMS based services. There are two cases can be considered:

Case 1:

Currently 22.261 [2] states that “A network slice can provide the functionality of a complete network, including radio access network functions and core network functions (e.g., potentially from different vendors). One network can support one or several network slices.”

However this doesn’t include IMS as part of the network slice.

The category of 5GC slice is different from IMS slice. 5GC slice types are categorized by the different types of network performance requirements (e.g. latency, bandwidth, etc.). But IMS slice type should be categorized by the different types of application requirements (e.g. MCPTT, Public Safety, etc.), which does not match to 5GC slice types.

Therefore it is possible that a user access to a specific IMS slice via different 5GC slices, e.g. a user subscribed to IMS video service may access to an IMS real time communication slice via eMBB slice and URLLC slice (e.g. video call when using AR/VR).

However, some applications may require the same network performance, so that a user can access different IMS slices via a single 5GC slice, e.g. a user may access via eMBB slice to IMS voice call slice or Public safety slice.

It is therefore proposed to consider IMS network slicing independent to 5GC slice as shown as the figure below.



Figure 4.6-1: IMS network slicing independent to 5GC slice

Case 2:

According to NGMN document - Description of Network Slicing Concept v1.0.8 [2] Clause 4.1, which is cited below, IMS can be a sub network instance of a network slice instance.

*Some examples of a network slice instance: Enhanced MBB, M2M, Enterprise and Industry etc.*

*Example of a sub network instance: IMS (IP Multimedia Subsystem) etc.*

*The concept is extensible to any scenario envisioned for an application of the network slice framework.*

So, another option is to include IMS as part of the E2E slice.

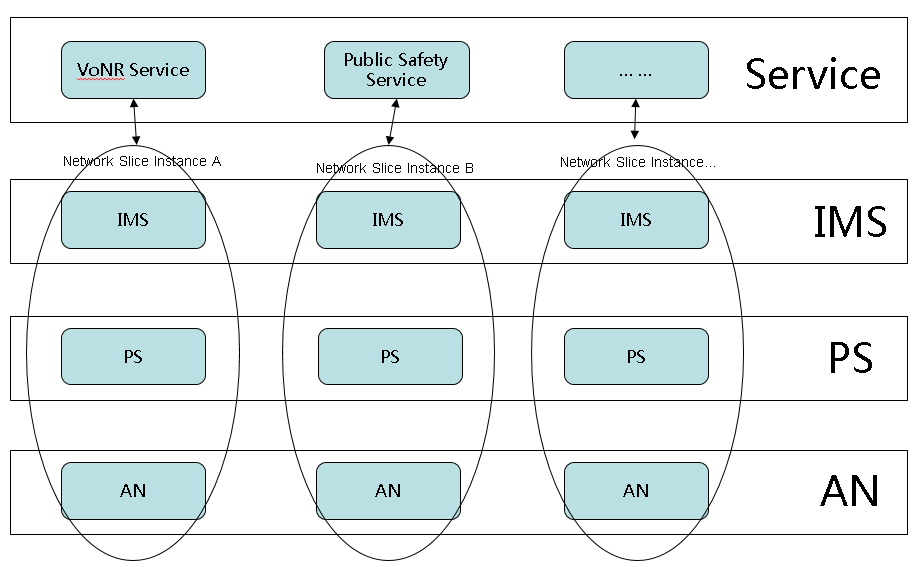


Figure 4.6-2: IMS network slicing as part of 5GC slice

### 4.6.2 Pre-conditions

NA

### 4.6.3 Service Flows

NA

### 4.6.4 Post-conditions

NA

### 4.6.5 Potential Impacts or Interactions with Existing Services/Features

None identified.

### 4.6.6 [Potential] Requirements

IMS network shall support the Network slicing requirements as defined in TS 22.261[2] clause 6.1.2.

Note: IMS network can either be part of network slice or an independent slice.

IMS network shall allow user to access a specific IMS slice via different 5GC slices.

IMS network shall allow users to access different IMS slices via a single 5GC slice.

## 4.7 Service visibility and usage

### 4.7.1 Description

As new real time services are emerging, more types of devices are emerging, including AR/VR glasses, cameras, robots, smart wearable, etc. So, UE will need to support many real time services.

Take remote guidance application as an example, where one user can use video communication to guide another user to perform a maintenance task.

### 4.7.2 Pre-conditions

The operator’s network supports remote guidance service based on video communication.

### 4.7.3 Service Flows

User A and B register with IMS network and get IMS network service list that includes the remote guidance service.

The service list is displayed on the user A and B UEs.

User B selects the remote guidance service, and connects to user A.

User A instructs user B with the maintenance task.

### 4.7.4 Post-conditions

User B completes repairs task successfully and ends the call/session.

### 4.7.5 Potential Impacts or Interactions with Existing Services/Features

As new services are developed operators want to advertise these services to the users so they can start using them. Hence, the IMS needs to provide a capability to indicate / notify the available service offering to the users and any changes to the service offering.

### 4.7.6 [Potential] Requirements

The IMS network shall provide a capability to indicate the available services offering, and changes to the available services..

## 4.8 Services using Artificial Intelligent (voiceprint) in real time communication

### 4.8.1 Description

With the development of cloud computing, improvement in CPU, GPU, and TPU computing capability, and development of various algorithms (Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), etc.), Artificial Intelligent (AI) application have made much progress in recent years. In the content awareness of service areas, such as image recognition, speech recognition and voiceprint recognition, there are mature applications in the markets of security, intelligent traffic system, and biometric authentication.

In additional, technologies of machine cognition and reasoning, such as Natural Language Processing (NLP), have been rapidly developing. There are a lots of CHATBOT (computer programs that mimic conversation with people using artificial intelligence.) and intelligently speakers (machine to copy intelligent human behavior.), representing voice control functions in the market.

These techniques can be used in real-time communication in public domain and the vertical industries to provide a more intelligent and convenient means of communication for users.

Hence, based on voiceprint, IMS can enable end users with real time authentication/validation of user information as given in the use case below.

### 4.8.2 Pre-conditions

User A’s voiceprint characteristic information is pre-provisioned in the IMS.

### 4.8.3 Service Flows

User A calls user B.

During the call user A asks B about user C contact information.

Before providing C’s contact information, B uses an IMS service to authenticate/validate the user A’s identity using A’s voiceprint, which user A agreed to have done.

User B gets the result of user A’s identity validation.

### 4.8.4 Post-conditions

Based on the validation result B takes the appropriate action (e.g. provide C information if A validation was successful).

### 4.8.5 Potential Impacts or Interactions with Existing Services/Features

None identified.

### 4.8.6 [Potential] Requirements

The IMS network shall be able to use the voiceprint analysis to authenticate the user.

The IMS network shall provide mechanisms to record the user permission for using voiceprint analysis.

The IMS network shall provide mechanisms to protect users' data related to voiceprint analysis.

## 4.9 Services using Artificial Intelligent ( on demand translation) in real time communication

### 4.9.1 Description

3GPP has a speech recognition framework (SRF) since release 6 [4]. Although services in this field have not been widely used, this previous study and the function requirement illustrate a use case for IMS to support automated voice services on the server side [5]:



Figure 4.9-1: On demand translation in real time communication

Some of popular scenarios for speech-to-speech translation are in the area of travel, customer support and online meeting. These scenarios are relevant to real time communication.

### 4.9.2 Pre-conditions

In an ongoing IMS call session, a foreign tourist User A is talking with a local person User B. User A don’t understand User B’s language.

User A is provisioned with the real time translation service in IMS network.

IMS is able to provide real time translation service on demand during the call session.

### 4.9.3 Service Flows

User A invokes the real time translation service on an active call which redirect B’s media (audio) to the translation server.

User A receive translated information of User B’s speech.

Based on operator’s policy and regional regulation for privacy, User B may be informed that User A is using a translation service.

### 4.9.4 Post-conditions

None.

### 4.9.5 Potential Impacts or Interactions with Existing Services/Features

None identified.

### 4.9.6 [Potential] Requirements

The IMS network shall be able to support real time translation service that can be invoked by a user on demand during a call and which allows the incoming media (e.g. audio) to be translated.

The IMS network shall be able to route the media (e.g. audio) that need to be translated to the translation server.

NOTE 1: The output from the translation server can be audio/text.

NOTE 2: Based on operator’s policy and regional privacy regulation, User B may be informed that User A is using a translation service.

NOTE 3: The use of real time translation service for emergency call should be considered and is subject to regional regulation.

# 5 Considerations

## 5.1 Considerations on security

For Artificial Intelligent (voiceprint) in real time communication service in clause 4.8, security measures to protect users’ data related to voiceprint analysis are required.

## 5.2 Considerations on Charing Aspect

For one source to many destinations communication service in clause 4.1, the operator may choose to distinguish the communication mode for charging (e.g. interactive or receive only) per operator’s policy. The communication mode (e.g. interactive, receive only, or switching among these modes) shall be reported in the charging information.

# 6 Conclusion and Recommendations

RTC is one of the key service in current mobile telecommunication services. As IMS supports RTC service, it is essential to consider how to enhance IMS for the emerging new RTC services.

In this study a number of use cases were analysed for enhancements to IMS including one to many asymmetric communication, UE communication with non-sip devices, delivery of information to control functions of the non-sip devices, fast IMS call setup time for Group Communication, enhancements on voice & video communications for supporting AR/VR, network slicing for IMS, service visibility and usage, services using AI (voiceprint) in RTC.

The study has resulted in a set of potential requirements as captured in the previous clauses.

It is therefore recommended that the potential requirements identified in the present TR be considered for the development of normative requirements.

Annex A:  
Change history

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Change history | | | | | | | | | | | |
| **TSG SA#** | **SA Doc.** | **SA1 Doc** | **Spec** | **CR** | **Rev** | **Rel** | **Cat** | **Subject/Comment** | **Old** | **New** | **WI** |
| [SP-80](http://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3373) | SP-180336 | [S1-181722](http://www.3gpp.org/ftp/tsg_sa/WG1_Serv/TSGS1_82_Dubrovnik/Docs/S1-181722.zip) | [22.823](http://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3373) | 0001 | 3 | [Rel-16](http://portal.3gpp.org/desktopmodules/Release/ReleaseDetails.aspx?releaseId=191) | B | Services using AI (instant translation) in real time communication | 16.0.0 | 16.1.0 | [FS\_enIMS](http://portal.3gpp.org/desktopmodules/WorkItem/WorkItemDetails.aspx?workitemId=770003) |