**Source: Rapporteur FS\_ULBC**

**Title: Permanent Document FS\_ULBC**

**Version: 0.1.1**

**Agenda item: 7.9**

**Document for: Agreement**

**Revision history:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Version** | **Date** | **Meeting** | **TDOC** | **Subject/Comment** |
| 0.1.0 | 2025-05-23 | SA4#132 | S4-251151 | Initial draft |
| 0.1.1 | 2025-05-23 | SA4#132 | S4-251154 | Update Clause 5.2 to include the methodology of obtaining channel characteristics |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

# 1 Introduction

The present document compiles candidate changes, open issues, incomplete text, dependencies from other group, action items, and considered timeline to 3GPP TR 26.940 “Study on Ultra Low Bitrate Speech Codecs”. TR 26.940 aims for developing recommendations for potential normative work on an ultra-low bit rate codec for voice over Geostationary Orbit (GEO) satellites.

The following clauses and subclauses are structured according to the objectives that are in scope of the FS\_ULBC SID [1]:

1. Document the application scenarios for ultra-low bit rate communication services taking into account the use cases and potential requirements documented in TR 22.887 related to IMS Voice Call Using GEO Access.

2. Study GEO channel characteristics and derive service-related dependencies, e.g. bitrates, mouth-to-ear delay or loss/delay/jitter profiles.

NOTE: Any impact of ultra-low bitrate voice codec in NB-IoT services is outside of the scope of the study and is expected to be addressed by other working groups.

3. Identify the relevant design constraints for such a codec, in coordination with other WGs, including

- Bit rates

- Sample rate and audio bandwidth

- Frame length

- Complexity and memory demands

- Algorithmic delay

- Packet loss concealment (PLC)

- Potential use of noise suppression as part of the codec

- Discontinuous transmission including voice activity detection and comfort noise

- Speech quality

- Robustness to non-speech input

4. Provide some evidence that the design criteria can be met, for example existing reference codecs.

5. Define performance requirements and identify appropriate test methodologies, regarding speech quality, intelligibility, conversational quality, in particular taking into account

a) Clean speech and noisy speech

b) Tandeming with existing IMS voice codecs

c) Clean channel and GEO channel conditions

6. Identify or develop objective measures to verify the design constraints as necessary (e.g., to measure complexity and memory demands)

7. Identify relevant reference codecs for comparison and evaluation purposes.

8. Coordinate work with other 3GPP groups e.g. SA2, RAN, CT1, and others as needed.

9. Define potential normative work item objectives and timeline.

This working procedure of TR and p-doc includes:

- Maintain one TR and one p-doc (this document)

- All contributions to the TR are expected to be submitted using pCRs

- Both pCRs and discussion papers may be be used to contributed to the p-doc.

- Brackets should be avoided when possible, and when used:

- Restricted to values only

- Never applied to complete text blocks

- Open issues in the TR are to be documented in the p-doc, for example the prioritization of application scenarios and related technical assumptions.

- The p-doc should keep track of the status of the individual study item objectives.

# 2 References

[1] SP-250378, "SID on Ultra Low Bitrate Speech Codec", China Mobile Com. Corporation, vivo, Fraunhofer IIS, Qualcomm Incorporated, Spreadtrum, Dolby Laboratories Inc., Xiaomi, Huawei, 2025.

[26132] 3GPP TS 26.132: “Speech and video telephony terminal acoustic test specification”.

[38811] 3GPP TR 38.811, “Study on New Radio (NR) to support non-terrestrial networks”.

[38821] 3GPP TR 38.821, “Solutions for NR to Non-Terrestrial Networks (NTN)”

[36321] TR 36.321, "Evolved Universal Terrestrial Radio Access (E-UTRA); Medium Access Control (MAC) protocol specification"

# 3 Definitions of terms, symbols and abbreviations

## 3.1 Terms

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

**example:** text used to clarify abstract rules by applying them literally.

## 3.2 Symbols

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

<symbol> <Explanation>

## 3.3 Abbreviations

For the purposes of the present document the following apply

# 4 Application scenarios

Editor’s Note:

1. Document the application scenarios for ultra-low bit rate communication services taking into account the use cases and potential requirements documented in TR 22.887 related to IMS Voice Call Using GEO Access.
2. Additional study areas or use cases, such as IMS voice call over NGSO or TN should be added with lower priority if time permits and once the exact requirements can be given.

## 4.1 Scenario 1: IMS Voice Call over GEO

### 4.1.1 Extracted technical assumptions and open issues

#### 4.1.1.1 Assumptions for Main Scenario

The following assumptions apply for Main Scenario described in clause 4.2.2.2.

- For the connection “UE1 – GEO satellite – Ground station” (UE1 uplink), the transmission data rate is significantly limited ([1-3] kbit/s), requiring an ultra-low bit rate codec fitting the transmission data rate for this link.

- For the connection “Ground station – GEO satellite – UE1” (UE1 downlink), the transmission data rate is expected to be limited similarly to UE1 uplink.

- For both uplink and downlink of UE1 it is expected that the link is subject to transmission errors reflecting GEO satellite access

Editor’s Note: "expected" should be replaced by more technical evidence when available (e.g., after coordination with RAN groups).

- The delay in uplink and downlink of UE1 is expected to be greater than the one of typical terrestrial networks.

Editor’s Note: "expected" should be replaced by more technical evidence when available (e.g., after coordination with RAN groups)

- For the connection "Core Network – UE2" (UE2 downlink), the transmission data rate of a regular TN network is available. This link could be covered either by an existing IMS codec (transcoding necessary) or by the same ultra-low bit rate codec as used for the satellite link (transcoding-free).

- To ensure seamless communication across different network types, roaming, etc. transcoding functionality in core network is likely needed.

Editor’s Note: More details may be added.

#### 4.1.1.2 Assumptions for Sub-Scenario

The following assumptions apply for Sub-Scenario described in clause 4.2.2.3.

- For both connections "UE1 – GEO satellite – Ground station" and "Ground station – GEO satellite – UE2" the transmission data rate is significantly limited ([1-3] kbit/s), requiring an ultra-low bit rate codec fitting this transmission data rate for these links.

- This scenario may allow both transcoded (ULBC 🡨🡪 existing IMS speech codecs 🡨🡪ULBC) and transcoding-free operation (ULBC end-to-end)

Editor’s Note: More details may be added.

## X Scenario X: TBD

### 4.X.1 Background

### 4.X.2 Scenario Description

### 4.X.3 Extracted technical assumptions and open questions

# 5 Channel characteristics and service-related dependencies

Editor’s Note:

Study GEO channel characteristics and derive service-related dependencies, e.g. bitrates, mouth-to-ear delay or loss/delay/jitter profiles as priority.

## 5.1 Architectural components and interfaces

### 5.1.1 Scenario 1: IMS Voice Call over GEO

### 5.1.X Scenario X:TBD

## 5.2 Channel characteristics

Editor’s Note:

- Study bitrates and loss/delay/jitter profiles.

### 5.2.1 Scenario 1: IMS Voice Call over GEO

#### 5.2.1.1 Introduction

This clause introduces the methodology of obtaining channel characteristics and results for developing design constraints and performance requirements for a codec supporting the main scenario as documented in Clause 4.2.1: IMS Voice Call over GEO.

#### 5.2.1.2 Delay error profiles

The delay-error profile is a model used to describe the network impairments—particularly delay and packet loss—that can impact real-time conversational services such as IMS voice call. Such profile typically reveals the GEO satellite channel characteristics and will be used to evaluate codec robustness, guide jitter buffer design and ensure a fair and comparable testing.

#### 5.2.1.3 End to end simulation model to derive delay error profiles

The intention of this methodology is to reuse the simulation model defined in Annex E of TS 26.132 [26132] to produce the delay error profile.

This Annex E reference LTE access scenario is illustrated in Figure 5.2.1.3-1. Building on the main scenario defined in Clause 4.2.1, the corresponding end-to-end GEO access scenario is shown in Figure 5.2.1.3-2. The primary distinction between the reference LTE scenario and the GEO voice main scenario lies in the introduction of the “new GEO channel”.



Fig.5.2.1.3-1: End-to-end channel of VoLTE using LTE access



Fig.5.2.1.3-2: End-to-end channel of main scenario for IMS voice call using NB-IoT (GEO) satellite access

Based on the functional description in Table E.1 of TS 26.132, the following input parameters are required to implement the simulation model:

**BLER\_tx / BLER\_rx**:

These parameters are required to simulate block error rates in both uplink and downlink.

NOTE: the resulted error trace based on Clause 5.2.2 will be used to serve as the BLER\_tx/BLER\_rx.

**[max\_tx / max\_rx**:

These define the maximum number of HARQ retransmissions for uplink and downlink respectively, which fall under RAN2 scope. In current specifications, NB-IoT supports at most two HARQ processes, which face constraints in high-latency GEO satellite scenarios. For IMS voice over GEO, HARQ feedback is suggested to be disabled per the standard of Release 18 [5].]

**drx\_cycle\_length**:

This parameter represents the duration of the DRX (Discontinuous Reception) cycle in milliseconds. It determines how frequently the device wakes up to monitoring possible scheduling grant. This parameter affects packet scheduling and transmission timing in the simulation context Annex E of TS 26.132. In addition, the values for LTE are 20-40ms, whether these values are suitable for GEO scenarios should be confirmed with RAN2.

**mis\_eNB1\_eNB2**:

This parameter represents the scheduling time mis-align between the two eNBs. In GEO scenarios, it indicates how long packets wait in the buffer before the next transmission opportunity. This should be determined primarily by RAN2 (responsible for dynamic scheduling or Semi-Persistent Scheduling) with possible input from RAN1 about physical layer timing relationship aspects.

**[max\_net\_delay / min\_net\_delay**:

These represent the delay range between eNB1 and eNB2. For GEO voice, they are considered similar to the LTE scenario, and legacy parameter values can be reused.]

Editor’s NOTE: whether the model for the delay between eNB1 and eNB2 for LTE scenarios well reflects the delay in deployment is FFS.

**nFrames**:

This refers to the number of frames for the simulation. In the reference LTE scenario, one IP packet corresponds to 20 ms of speech. In contrast, the GEO voice scenario introduces additional considerations shown as follows due to the propagation delay from GEO satellite altitude.

* **Speech sequence (frame length)**: For GEO, a longer frame length may be used. The maximum frame length of **80 ms**, as defined by 3GPP, is assumed in this simulation. Final confirmation is expected from SA4.
* **Voice packet size**: This depends on the protocol overhead as illustrated in Figure 5.2.1.1-3 for the reference LTE access scenario and Figure 5.2.1.1-4 for the GEO voice main scenario. The exact overhead depends on the transport path of the RTP packet—user plane or control plane, via IP or via Non-IP (NIDD)—and must be confirmed by RAN2 and SA2. For simulation purpose, the assumed range is as follows:

- **Lower bound**: ~**7 bytes**. Assuming user plane with ROHC (3 bytes) and NB-IoT protocol layers: PDCP (1 byte), RLC (1 byte), MAC (2 bytes);

- **Upper bound**: ~**52 bytes**. Assuming control plane with IPv4/UDP/RTP (20/8/12 = 40 bytes), NAS reduction (7 bytes, optimized by CT1), and NB-IoT layers: RRC (2 bytes), RLC (2 bytes), MAC (1 byte);



Fig. 5.2.1.3-3: VoIP RTP packet in reference LTE access scenario





Fig.5.2.1.3-4: Example of RTP packet in GEO voice main scenario

- **RTP Payload Size**: This is computed as the product of frame length and codec bit rate.

Editor’s Note: whether the size of RTP payload affects the delay-error profile is FFS.

Once the parameters regarding GEO channel are confirmed, the simulation methodology as described in Table E.1 will be updated with these new parameters and used to produce the required delay-error profiles.

### 5.2.2 Simulation Model to derive error traces

The NTN link consists of a service link (between the UE and the satellite) and a feeder link (between the satellite and the ground station). The bottleneck is the link due to the limited TX power and small antenna at the UE. The feeder link is typically characterized by large capacity and high reliability and can be abstracted as an ideal link in the end-to-end simulation. The RAN simulation addresses the service link only.

The objective is to generate multiple loss traces for a combination of frame loss rate (target BLER), raw bitrate (TBS), voice bundling period and Doppler spread, while maintaining channel consistency among different combinations.

The multiple loss traces are the result of using multiple random seeds, and the number is 10. For each combination, all 10 seeds are used in generating the error traces.

Each trace represents a duration of 400 seconds (or 6.67 minutes). Therefore, for 80ms bundling, there are 5000 TBs, and for 160ms bundling there are 2500 TBs.

The following parameters are for the uplink of the service link.

**Channel model**: NTN-TDL-C [38811]

**Modulation:** QPSK

**SCS**: 3.75kHz, 15kHz

**Number of tones:** 1

**Voice bundling period**: 80ms, 160ms, 320ms

NOTE: the 40ms bundling is not considered because for SCS 3.75kHz the minimum time-domain allocation is 32ms and it leaves insufficient time for downlink data (NPDSCH) and control (NPDCCH) transmissions in the same 40ms time interval.

**Doppler spread**: 1Hz, 5 Hz

**Target BLER**: 1%, 2%, 6%, 10%

**Target SNR**: Company report, 3GPP SET-1 SNR +[X] dB, X is TBD, where 3GPP SET-1 parameters are defined in [38821] and the [X] dB accounts for better performance of commercial satellites.

**TBS** **values and PHY bitrates**: The TBS values are selected from table 16.5.1.2-2 for NB-IoT for NPUSCH in TS36.213 and the corresponding PHY bitrates and codec bitrate (assuming 8 bytes of packet header with RoHC) are calculated for each bundling period.

[

Table 1 TBS and PHY bitrate for 80ms bundling

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| TBS (bits) | 144 | 256 | 328 | 424 |
| PHY bitrate (kbps) | 1.8 | 3.2 | 4.1 | 5.3 |
| Codec bitrate (kbps) | 1.0 | 2.4 | 3.3 | 4.5 |

Table 2 TBS and PHY bitrate for 160ms bundling

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| TBS (bits) | [208] | 424 | 600 | 808 |
| PHY bitrate (kbps) | [1.30] | 2.65 | 3.85 | 5.05 |
| Codec bitrate (kbps) | [0.9] | 2.25 | 3.45 | 4.65 |

Table 3 TBS and PHY bitrate for 320ms bundling

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| TBS (bits) | [328] | 776 | 1000 | 1544 |
| PHY bitrate (kbps) | [1.025] | 2.425 | 3.125 | 4.825 |
| Codec bitrate (kbps) | [0.8025] | 2.225 | 2.925 | 4.625 |

]

Editor’s NOTE: Company can report candidate values of TBS.

NOTE: A PHY bitrate is considered feasible if the required SNR to meet a target BLER is lower than the target SNR.

**Channel consistency**: The same set of channel realizations are used across all combinations.

### 5.2.3 Results

Editor’s Note: the results are FFS

### 5.2.X Scenario X:TBD

## 5.3 service-related dependencies

Editor’s Note:

- Study mouth-to-ear delay.

### 5.3.1 Scenario 1: IMS Voice Call over GEO

### 5.3.X Scenario X:

# 6 Design constraints

Editor’s Note:

2.. Identify the relevant design constraints for such a codec, in coordination with other WGs, including:

- Bit rates

- Sample rate and audio bandwidth

- Frame length

- Complexity and memory demands

- Algorithmic delay

- Packet loss concealment (PLC)

- Potential use of noise suppression as part of the codec

- Discontinuous transmission including voice activity detection and comfort noise

- Speech quality

- Robustness to non-speech input

- Identify or develop objective measures to verify the design constraints as necessary (e.g., to measure complexity and memory demands)

## [6.1 The status of DCs

|  |  |  |  |
| --- | --- | --- | --- |
| **External Dependency** | **Information from External Groups** | **SA4's Assumption** | **Open Issues** |
| DC: Bitrates |  | *e.g., A preliminary set of bitrates assumed by SA4 includes xxx-xxkpbs.* |  |
| DC: DTX/CNG |  |  |  |
| DC: Frame length | *e.g., S4-XXXXX an LS from RAN that indicats that the frame length should...* |  |  |
| DC: PLC loss/dly/error proiles |  |  |  |
| DC: Alg. Delay |  |  |  |

|  |  |  |  |
| --- | --- | --- | --- |
| **Interdependency** | **Core Influencing Factors** | **Progress** | **Open Issue** |
| **DC: Bitrate** | Frame length (Major and External) |  |  |
| Robust Non-Speech |  |  |
| Evidence DCs |  |  |
| Noise Supression |  |  |
| Study GEO channel characteristics, derive service-related dependencies (Major and External) |  |  |
| **DC: Complexity, Memory** | Objective Measures (Major) |  |  |
| DC:Robust Non-Speech |  |  |
| DC: noise suppression |  |  |
| Evidence DCs |  |  |
| **DC: Sample rate, audio bandwidth** | Evidence DCs |  |  |

]

# 8 Existing technologies and feasibility evidence

Editor’s Note:

Provide some evidence that the design criteria can be met, for example existing reference codecs.

# 7 Performance requirements

Editor’s Note:

1.Define performance requirements and identify appropriate test methodologies, regarding speech quality, intelligibility, conversational quality, in particular taking into account:

- Clean speech and noisy speech

- Tandeming with existing IMS voice codecs

- Clean channel and GEO channel conditions

2.Identify relevant reference codecs for comparison and evaluation purposes.

## 7.1 The status of PRs.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Core Influencing Factors** | **Progress** | **Open Issue** |
| **Performance requirements/speech quality** | DC: Sample rate and audio bandwidth |  |  |
| DC: Bitrates (External) |  |  |
| DC: Frame length |  |  |
| DC: PLC (External) |  |  |
| DC: Algorithmic Delay |  |  |
| DC: Complexity, Memory |  |  |
| Test Methodolgies |  |  |
| DC:noise suppression |  |  |
| DC:DTX/CNG |  |  |
| DC:Robust Non-Speech |  |  |
| Evidence DCs |  |  |
| Reference codec |  |  |

## 7.2 Clean speech and noisy speech

## 7.3 Tandeming with existing IMS voice codecs

## 7.4 Clean channel and GEO channel conditions

# 9 Test methodologies

# 10Considered work plan for potential normative work

# 