3GPP TSG-SA WG4 Meeting #132S4-251125

Japan, Fukuoka, 19 – 23 May 2025

**Source: Fraunhofer IIS, vivo, Qualcomm Incorporated, Skylo**

**Title: Estimation mouth to ear delay for GEO NTN voice**

**Spec: 3GPP TR 26.940**

**Agenda item: 7.9**

**Document for: Agreement**

**1. Introduction**

At the recent SA meeting, the “Study on Ultra Low Bitrate Speech Codec” has been approved. According to the WID description [1], the study includes also service-related aspects such as mouth-to-ear delay for voice services over GEO.

**2. Reason for Change**

The present document provides an estimation of the missing mouth-to-ear delay for IMS voice call over GEO satellites.

.

**4. Proposal**

It is proposed to agree the following changes to 3GPP TR 26.940.

NOTE: This document assumes that the application scenarios in S4-250591 are agreed.

\* \* \* First Change \* \* \* \*

## Y Estimation of mouth to ear delay for GEO scenarios

### Y.1 Overview

This clause estimates the mouth to ear (M2E) delay for IMS voice call over GEO satellites based on the application scenario introduced in clause 4.2. Two sub-scenarios are considered:

**- Main Scenario (see clause 4.2.2.2):** UE1 is connected via satellite while UE2 is connected via terrestrial network which corresponds to the signal flow UE1 àGEO satellite àGround stationàCore networkà eNodeB àUE2

* **Sub-Scenario 1 (see clause 4.2.2.3):** Both UEs are connected to a GEO satellite which corresponds to the signal flow UE1 àGEO satellite àGround stationàCore networkàGround stationàGEO satelliteàUE2

This approach aims to estimate the maximum and minimum delay components in the signal flow and finally to estimate a range of the. mouth-to-ear delay accordingly. The estimation assumes jitter free case and no network congestion.

NOTE: In real deployments, other jitter and network conditions may occur.

Editor’s note: The scenarios and the terminology of this clause needs to be aligned with clause 4.) “Application Scenario” where a detailed description of the call scenarios is expected.

### Y.2 Delay components

#### Y.2.1 Overview

In this clause, the individual delay components that contribute to the mouth-to-ear delay are introduced and derived. The derived values are independent of the signal flow direction.

#### Y.2.2 UE Delay considering IMS codecs

[TS 26.131](https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=1408) [D.1] defines the internal UE delay requirements and objectives depending on the components codec (frame size and algorithmic delay), air interface, jitter buffer depth and vendor specific delay budget. The UE delays in sending (UE1) and receiving directions (UE2) are not separated in TS 26.131, however the sum of the sending and receiving delays can be considered together.

The jitter buffer delay budget contains 40ms if the packet duration is 20ms and it includes the expected jitter profiles for terrestrial network transmission. In case of 40ms packet duration, the budget is doubled to 80ms, which is not further discussed. The value for the air interface in [D.1] just reflects the delay between UE and measurement equipment (2 ms) and needs be replaced by the expected delay for real air interface, i.e. air interface to GEO satellite or terrestrial network.

For MTSI-based speech only services with LTE and NR, the UE delay is outlined in the following table:

Table Y.2.2-1 UE delay components

|  |  |  |
| --- | --- | --- |
|  |  UE delay in ms (Performance objective)(Note 2) | UE delay in ms (Maximum requirement)(Note 2) |
| Frame size (Note 1) | 20 | 20 |
| alg. Codec Delay (Note 1) | 5 | 12 |
| JBM (jitter free) (Note 3) | 40 | 40 |
| Vendor specific budget (Note 4)  | 83 | 123 |
| UE delay Ts+Tr | 148 | 195 |
| Note 1: Values reflect the IMS codecs AMR/AMR-WB/EVSNote 2: Requirements and Performance Objectives apply to the UE delay only (sum of send (Ts) and receive (Tr) delays) and only for MTSI-based speech-only with LTE, NR or WLAN access in error and jitter free conditions.Note 3: JBM delay is considered as constant independent of the frame size.Note 4: Vendor specific budget of TS 26.131 may change for GEO satellite connectivity |

Editor’s note: For GEO satellite access the radio air interface delay needs to be taken into account.

#### Y.2.3 Core network delay

The delay contribution of the core network consists of the packet transmission delay between two network entities, e.g. ground station to core network or core network to eNodeB. In case of the interop scenario GEO NTN to TN network, an additional delay component for transcoding needs to be considered. Assuming the frame size of both codecs is identical or a multiple of each other, only the algorithmic codec delay contributes to the transcoding delay, i.e. 5ms for AMR/AMR-WB or 12ms for EVS, and an additional delay margin for the processing of the transcoding (2 ms). This means, transcoding with AMR/AMR-WB adds 7ms and with EVS adds 14ms.

Table Y.2.3-1 Core network delay components

|  |  |  |
| --- | --- | --- |
|  | Minimum delay in ms | Maximum delay in ms |
| Network delay ground station to core network (Delay\_GSCN) | [ 5 Note1-1 , 20 Note1-2]  | [200 Note2] |
| Network delay eNodeB to core network (Delay\_eNBCN) | 5 | 20 |
| Transcoding | 7 | 14 |
| [Note1-1: In [D.2] 5 ms network latency is assumed][Note1-2: TS 23.501 assumes a static delay value for the CN PDB of 20ms between a UPF and 5G-AN. ][Note2: In some NTN deployments, the core network may need to be located far from the ground station due to factors like user distribution, geography, or other practical considerations. As a result, latency can increase, ping statistics between continents, for example, can reach up to 200ms.] |

#### Y.2.4 Transmission delay UE – GEO - Ground station

Clause 7.4.2 of [D.2] defines the KPI requirement for GEO based satellite access, i.e. 280ms. TR 36.763 clause 7.1.1 describes the max. and min. propagation delay contribution which depends on the location of the UE within the beam. As a result, the round-trip-delay can differ by 64ms which corresponds 32ms for one-way transmission. It is proposed to consider the 280ms as the max. transmission delay and consequently 248ms (280ms – 32ms) as the minimal transmission time. This assumes no retransmission over the GEO satellite link.

Table Y.2.4-1 Transmission delay GEO satellite

|  |  |  |
| --- | --- | --- |
|  | Minimum delay in ms | Maximum delay in ms |
| GEO transmission delay | 248 | 280 |
| Note: Transmission delay ground station to core network counted in Y.2.3-1. |

#### Y.2.5 ULBC Delay components

Table Y.2.2-1 lists the algorithmic delay for the IMS codecs AMR and EVS, i.e. in range of 5ms to 12ms. For ULBC, different delay values may result from codec processing delays as well as algorithmic delays. Exact numbers are for further study.

### Y.3 Estimation of Mouth-to-ear delay

Given the values in Y.2 the mouth-to-ear delay for scenario can be estimated for the two scenarios outlined in Y.1 by summing up the delay components according to the signal flow to derive a lower (minimum values as in tables Y.2.2-1, Y.2.3-1, Y.2.4-1) and an upper bound (maximum values as in tables Y.2.2-1, Y.2.3-1, Y.2.4-1).

As the bitrate for GEO satellite link is very restricted, options for minimizing the protocol overhead need to be considered. One option to reduce the protocol overhead are larger frame sizes or frame aggregation as the protocol stack is transmitted less often. Therefore, the following table outlines the delay values for codec frame sizes of 20ms, and in addition derived for 40ms, 80ms, 160ms and 320ms.

Editor’s Note: Current values assume algorithmic delay of AMR and EVS as given in Y.2.2-1. ULBC Delay components documented in Y.2.5 still need to be addressed. For the min. Delay\_GSCN, 20ms is assumed.

Table Y.3-1 Mouth-to-ear delay estimation depending on codec frame size

|  |  |  |
| --- | --- | --- |
| Frame size in ms | Mouth to ear delay main scenario in ms(GEO - TN) (Note 1) | Mouth to ear delay sub-scenario 1 in ms(GEO - GEO) (Note 2) |
| **lower bound** | **upper bound** | **lower bound** | **upper bound** |
| 20 | 428 | 712 | 684 | 1155 |
| 40 | 448 | 732 | 704 | 1175 |
| 80 | 488 | 772 | 744 | 1215 |
| 160 | 568 | 852 | 824 | 1295 |
| 320 | 728 | 1012 | 984 | 1455 |
| Note 1: UE(frame size) + UE(alg. Delay) + UE(JBM) + UE(Vendor)+GEO transmission+Delay\_GSCN+Delay\_eNBCNNote 2: UE(frame size) + UE(alg. Delay) + UE(JBM) + UE(Vendor)+2x GEO transmission+2x Delay\_GSCN |

Editor’s note: The scenarios and the terminology of this clause needs to be aligned with clause 4.) “Application Scenario” where a detailed description of the call scenarios is expected.

\* \* \* Next Change \* \* \* \*

[D.1] 3GPP [TS 26.131](https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=1408), „Terminal acoustic characteristics for telephony; Requirements“

[D.2] 3GPP [TS 22.261](https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3107), “Service requirements for the 5G system”

[D.3] 3GPP TS.36.763, “Study on Narrow-Band Internet of Things (NB-IoT) / enhanced Machine Type Communication (eMTC) support for Non-Terrestrial Networks (NTN)”

\* \* \* End of Changes \* \* \* \*

[1] Tdoc [SP-250378](https://www.3gpp.org/ftp/tsg_sa/TSG_SA/TSGS_107_Incheon_2025-03/Docs/SP-250378.zip), “Study on Ultra Low Bitrate Speech Codec”