**3GPP TSG SA WG 1 Meeting #104 S1-23xxxx**

**Chicago, USA, 13 - 17 November 2023** *(revision of S1-23xxxx)*

**Source: Huawei, Samsung**

**pCR Title: Pseudo-CR on addition of KPI requirements**

**Draft Spec: 3GPP TS 22.156 v1.0.0**

**Agenda item: 7.3.2**

**Document for: Approval**

**Contact: Alice Li ( alice.li1@huawei.com)**

*Abstract: This pCR proposes updates to the KPI requirements based on the agreements made in SA1 meeting #103.*

**1. Introduction**

As planned, the agreements made (to TR 22.856) in the last SA1 meeting need to be introduced as normative requirements in TS 22.156.

**2. Reason for Change**

The agreements captured in TR 22.856 CR0003 need to be introduced as the normative requirements in TS 22.156.

**3. Conclusions**

The main changes include:

- additional reference is added

- in table 6.2-1

- new rows are added

- additional notes are added

**4. Proposal**

It is proposed to agree the following changes to 3GPP TS 22.156 v1.0.0.

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

# 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 22.228: "Service requirements for the Internet Protocol (IP) Multimedia core network Subsystem (IMS)".

[3] ITU-T Recommendation Y.3090 (02/22): "Digital twin network - Requirements and architecture" (https://www.itu.int/rec/T-REC-Y.3090-202202-I).

[4] 3GPP TS 22.101: "Service principles".

[5] ITU-T Recommendation F.703 (11/00): "Multimedia conversational services".

[6] European Commission: "Shaping Europe's digital future", <https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/shaping-europe-digital-future_en>.

[7] 3GPP TS 22.261: "Service requirements for the 5G system".

[8] 5GAA: "C-V2X Use Cases Volume II: Examples and Service Level Requirements", 5G Automobile Association White Paper, https://5gaa.org/wp-content/uploads/2020/10/5GAA\_White-Paper\_C-V2X-Use-Cases-Volume-II.pdf <accessed 02.09.22>.

[9] O. Holland et al., "The IEEE 1918.1 "Tactile Internet" Standards Working Group and its Standards," Proceedings of the IEEE, vol. 107, no. 2, Feb. 2019."

[10] A. Ebrahimzadeh, M. Maier and R. H. Glitho, "Trace-Driven Haptic Traffic Characterization for Tactile Internet Performance Evaluation," 2021 International Conference on Engineering and Emerging Technologies (ICEET), 2021, pp. 1-6.

[11] glTF 2.0 specification, <https://registry.khronos.org/glTF/specs/2.0/glTF-2.0.html>, accessed 02/02/2023.

\* \* \* Second Change \* \* \* \*

6.2 Performance requirements

The 5G system shall support various mobile metaverse services with the following KPIs.

NOTE: Unless stated otherwise, the "Max allowed end-to-end latency" refers to the maximum transmission delay expected between a UE and the mobile metaverse server or vice-versa.

Table 6.2-1: Performance requirements for Mobile Metaverse Services

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Use Cases | Characteristic parameter (KPI) | | | | Influence quantity | | | | | Remarks |
| Max allowed end-to-end latency | Service bit rate: user-experienced data rate | Reliability | Area Traffic capacity | Message size (byte) | Transfer Interval | Positioning accuracy | UE Speed | Service Area |
| 5G-enabled Traffic Flow Simulation and Situational Awareness  (NOTE 2) | [5-20] ms (NOTE 1) | [10~100] Mbit/s  [8]  (NOTE 6) | > 99.9% | [TBD] Tbit/s/km2  (NOTE 5) | - | 20~100 ms  (NOTE 3) | - | < 250 km/h | City or Country wide  (NOTE 4) | UL |
| Collaborative and concurrent engineering | [≤10] ms  [9]  (NOTE 7) | [1-100] Mbit/s  [9] | [> 99.9%]  [9] | [1.55] Tbit/s/km2  (NOTE 8) | Video: 1500  Audio: 100  [9] | - | - | Stationary or Pedestrian | typically  < 100 km2  (NOTE 9) | UL and DL audio/video |
| [5] ms UL  [1-50] ms DL  [9]  (NOTE 7) | [<1] Mbit/s  [9] | [> 99.9%] (without compression)  [> 99.999%] (with compression (NOTE 10))  [10] | [2.25] Tbit/s/km2  (NOTE 8) | 1 DoF: 2-8  3 DoFs: 6-24  6 DoFs: 12-48  [9] | 0.25-10 ms  [9] |  |  |  | UL and DL haptic feedback |
| Metaverse-based Tele-Operated Driving  (NOTE 16) | [100] ms [8] (NOTE 11) | [10~50] Mbit/s [8] | 99% [8] | [~360] Mbit/s/km2  (NOTE 14) | - | 20~100 ms [8]  (NOTE 12) | [10] cm [8] | [10-50] km/h (vehicle) [8]  Stationary or Pedestrian | Up to 10 km radius [8]  (NOTE 13) | UL real-time vehicle data (video streaming and/or sensor data) [8] |
| [20] ms [8] | [0.1~0.4] Mbit/s [8] | 99.999% [8] | [~4] Mbit/s/km2  (NOTE 14) | Up to 8Kb  [8] | 20 ms [8]  (NOTE 12) | [10] cm [8] | [10-50] km/h (vehicle) [8]  Stationary or Pedestrian | Up to 10 km radius [8]  (NOTE 13) | DL control traffic (commands from the remote driver) [8]. |
| 1-20 ms  (NOTE 15) | 16 kbit/s -2 Mbit/s  (without haptic compression encoding);  0.8 - 200 kbit/s  (with haptic compression encoding)  (NOTE 15) | 99.999%  (NOTE 15) | [~20] Mbit/s/km2  (NOTE 14) | 2-8 (1 DoF) (NOTE 15) |  |  | Stationary or Pedestrian | Up to 10 km radius [8]  (NOTE 13) | Haptic feedback |
| Viewports streaming from rendering device to AR glasses through direct device connection (tethered/relaying case) (NOTE 17) | 10 ms (i.e., UL+DL between AR Glasses display and the rendering UE) (NOTE 18) | [200-2000] Mbit/s | 99.9 %  (NOTE 18) | - | - | - | - | Stationary or pedestrian (between rendering device and AR glasses) | Up to direct device connection ranging | Immersive AR interactive experience: tethered link |
| Pose information from AR glasses to rendering device through direct device connection (tethered/relaying case) (NOTE 17) | 5 ms  (NOTE 18) | [100-400] Kbit/s  (NOTE 18) | 99.9 %  (NOTE 18) | - | - | - | - | Stationary or pedestrian (between rendering device and AR glasses) | Up to direct device connection ranging |
| Movie streaming from metaverse server to the rendering device (NOTE 20) | Only relevant for live streaming.  [1-5] s in case of live streaming | [0.1-50] Mbit/s (i.e., covering a complete OTT ladder from low resolution to 3D-8K) (NOTE 19) | 99.9 % | - | - | - | - | [up to 500 km/h] | - | Immersive AR interactive experience: NG-RAN multimodal communication link |
| Avatar information streaming between remote UEs (end to end) | 10 ms (i.e., 20ms between both UEs excluding metaverse server processing time) (NOTE 22) | [0.1-30] Mbit/s  (NOTE 21) | 99.9 % | - | - | - | - | [up to 500 km/h] | - |
| Interactive data exchange: voice and text between remote UEs (end to end) (NOTE 22) | 10 ms (i.e., 20ms between both UEs excluding metaverse server processing time) | [0.1-0.5] Mbit/s | 99.9 % | - | - | - | - | [up to 500 km/h] | - |
| NOTE 1: The mobile metaverse server receives the data from various sensors, performs data processing, rendering and provide feedback to the vehicles and users.  NOTE 2: Examples of typical data volume including 1) camera: 10 Mbit/s per sensor (unstructured), 2) LiDAR: 90 Mbit/s per sensor (unstructured), 3) radar: 10 Mbit/s per sensor (unstructured), and 4) real-time Status information including Telemetry data: [< 50 kbit/s] per sensor/vehicle/VRU (structured). This is to support at least 80 vehicles and 1600 users present at the same location (e.g. in an area of 40m\*250m) to actively enjoy immersive metaverse services for traffic simulation and traffic awareness, the area traffic capacity is calculated considering 2 cameras, 2 Radars, 2 LiDARs on road side, 1600 user’s smart phones and 80 vehicles with 7 cameras, 4 radar and 2 LiDAR for each vehicle.  NOTE 3: The frequency considers different sensor types such as Radar/LiDAR (10Hz) and camera (10~50Hz).  NOTE 4: The service area for traffic flow simulation and situational awareness depends on the actual deployment, for example, it can be deployed for a city or a district within a city or even countrywide. In some cases a local approach (e.g. the application servers are hosted at the network edge) is preferred in order to satisfy the requirements of low latency and high reliability.  NOTE 5: The calculation is this table is done per one 5G network, in case of N 5G networks to be involved for such use case in the same area, this value can be divided by N.  NOTE 6: User experienced data rate refers to the data rate needed for the vehicle or human, the value is observed from industrial practice.  NOTE 7: The network based conference focus is assumed, which receives data from all the participants, performs rendering (image synthesis), and then distributes the results to all participants. As rendering and hardware introduce some delay, the communication delay for haptic feedback is typically less than 5ms.  NOTE 8: To support at least 15 users present at the same location (e.g. in an area of 20m\*20m) to actively enjoy immersive Metaverse service concurrently, the area traffic capacity is calculated considering per user consuming non-haptic XR media (e.g. for video per stream up to 40000 kbit/s) and concurrently 60 haptic sensors (per haptic sensor generates data up to 1024 kbit/s).  NOTE 9: In practice, the service area depends on the actual deployment. In some cases a local approach (e.g. the application servers are hosted at the network edge) is preferred in order to satisfy the requirements of low latency and high reliability.  NOTE 10: The arrival interval of compressed haptic data usually follows some statistical distributions, such as generalized Pareto distribution, and Exponential distribution [10].  NOTE 11: The end-to-end latency does not include sensor acquisition or actuator control on the vehicle side, processing, and rendering on the user side (estimated additional 100ms total). Target e2e user experienced max delay depends on reaction time of the remote driver (e.g. at 50km/h, 20ms means 27cm of remote vehicle movement).  NOTE 12: UL data transfer interval around 20ms (video) to 100ms (sensor), DL data transfer interval (commands) around 20ms.  NOTE 13: The service area for teleoperation depends on the actual deployment; for example, it can be deployed for a warehouse, a factory, a transportation hub (seaport, airport etc.), or even a city district or city. In some cases, a local approach (e.g., the application servers are hosted at the network edge) is preferred to satisfy low latency and high-reliability requirements.  NOTE 14: The area traffic capacity is calculated for one 5G network, considering 4 cameras + sensors on each vehicle. Density is estimated to 10 vehicles/km2, each of the vehicles with one user controlling them. [8]  NOTE 15: KPI comes from [7] clause 7.11 “remote control robot” use case.  NOTE 16: Examples of typical data volume including 1) ~8Mbps video stream. Four cameras per vehicle (one for each side): 4\*8=32Mbps. 2) sensor data (interpreted objects), assuming 1 kB/object/100 ms and 50 objects: 4 Mbps [8].  NOTE 17: These KPIs are only valid for cases where the viewport rendering is done in the tethered device and streamed down to the AR glasses. In the case of rendering capable AR glasses, these KPIs are not valid.  NOTE 18: These values are aligned with the tactile and multi-modal communication KPI table in TS 22.261 [7], clause 7.11.  NOTE 19: These values are aligned with “high-speed train” DL KPI from TS 22.261 [7] cl 7.1  NOTE 20: To leverage existing streaming assets and delivery ecosystem, it is assumed that the legacy streaming data are delivered to the rendering device, which incrusts this in the virtual screen prior to rendering. For a live streaming event, the user-experience end-to-end latency is expected to be competitive with traditional live TV services, typically [1-5] seconds.  NOTE 21: For example, the glTF format [11] can be used to deliver avatar representation and animation metadata in a standardized manner. Based on this format, the required bitrate for transmitting such data is highly dependent on avatar’s complexity (e.g., basic model versus photorealistic).  NOTE 22: These values are aligned with “immersive multi-modal VR” KPIs in TS 22.261 [7], clause 7.11. | | | | | | | | | | |