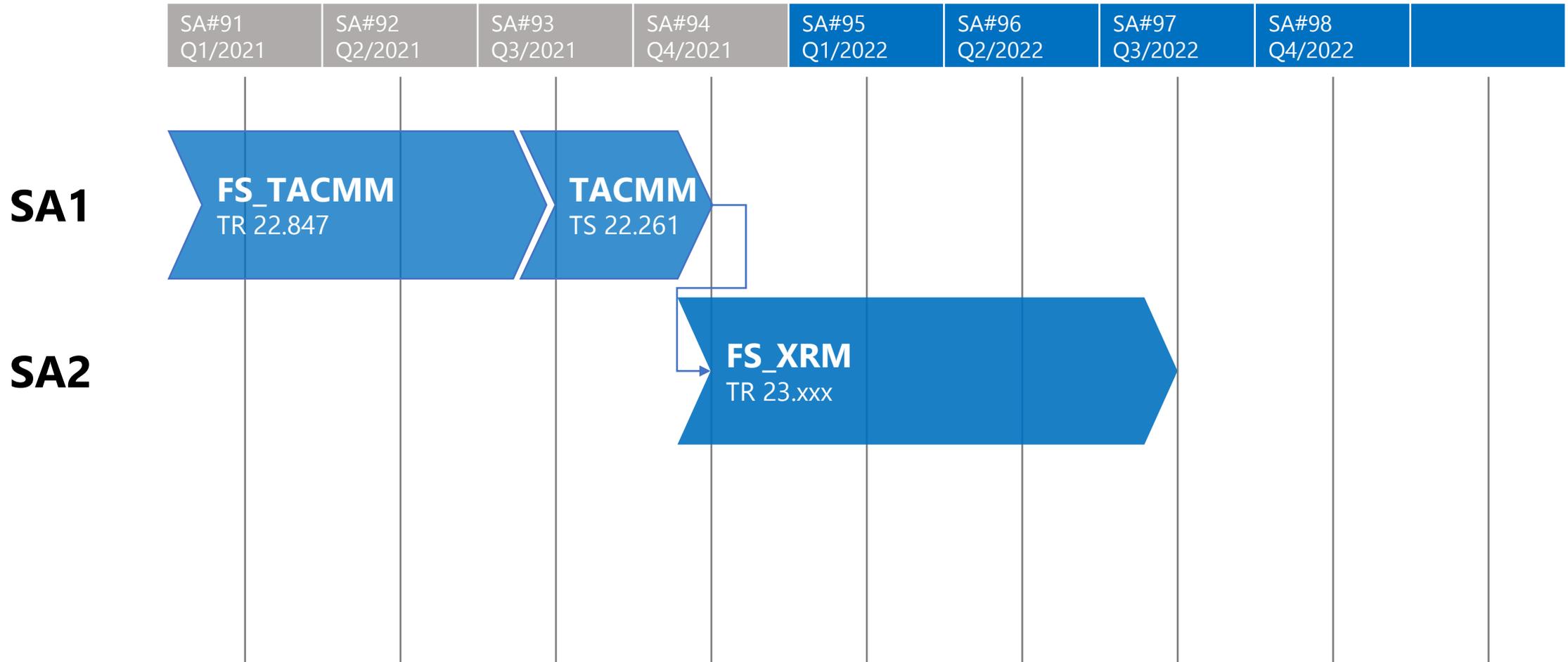


Brief Summary on Tactile and Multi-Modality Communication Services (FS_TACMM, TACMM) in SA1

Samsung Electronics Co., Ltd, Dolby Laboratories Inc.

Introduction



Tactile And Multi-Modality Communication Services

- SA1 Study (FS_TACMM: TR 22.847) '21.03→'21.09 (100%)
- SA1 Stage 1 (TACMM: CR on TS 22.261) '21.09→'21.12 (100%)

- TR Use case
 - #1 Immersive multi-modal Virtual Reality (VR) application
 - #2 Remote control robot
 - #3 Immersive VR games
 - #4 Support of skillset sharing for cooperative perception and maneuvering of robots
 - #5 Haptic feedback for a personal exclusion zone in dangerous remote environments
 - #6 Live Event Selective Immersion
 - #7 Support for IEEE P1918.1 architecture
 - #8 Virtual factory

UC #1 Immersive multi-modal Virtual Reality (VR) application (1 of 3)

- The application user utilizes the devices to experience immersive multi-modal VR application. The user powers on the devices to connect to the application server, then the user starts the gaming application.
- During the gaming running period, **the devices periodically send the sensing information to the application server, including: haptic and/or kinesthetic feedback signal information** which is generated by haptic device, and the sensing information such as positioning and view information which is generated by the VR glasses.
 - NOTE 1: **The devices may send the haptic data and the sensing data with different periodic time.** As an example, the device may send one packet containing haptic information to the application server every 2ms, and send the packets related to sensing information to application server every 4ms. **Thus the haptic data and sensing data may be transferred in 5G network via two separate flows.**
- According to the uplink data from the devices, the **application server** performs necessary process operations on immersive game reality including **rendering and coding the video, the audio and haptic model data**, then **application server periodically sends the downlink data to the devices, with different time periods respectively, via 5G network.**
- The devices, respectively, receive the data from the application server and present the related sensing including video, audio and haptic to the user.
 - NOTE 3: To obtain more realistic and compelling virtual environments, **network assistance may be needed to ensure synchronisation thresholds between different modal data thus improve the end users' sense of presence and realism.**



UC #1 Immersive multi-modal Virtual Reality (VR) application (2 of 3)

- [PR 5.1.6-1] The 5G System shall provide the network connection to address the [KPIs for immersive multi-modal VR applications](#), see table 5.1.6-1.
- [PR 5.1.6-2] Due to the separate handling of the multiple media components, synchronization between different media components is critical in order to avoid having a negative impact on the user experience (i.e. viewers detecting lack of synchronization). Applying synchronization thresholds in the 5G system may be helpful in support of immersive multimodal VR applications when the [synchronization threshold between two or more modalities is less than the latency KPI](#) for the application.
- [PR 5.1.6-3] The 5G network shall support a mechanism [to allow an authorized 3rd party to provide QoS policy for multiple flows](#) (e.g., haptic, audio and video) of multiple UEs associated with a multi-modal application. The policy may contain e.g. coordination information.
- [PR 5.1.6-4] The 5G system shall support a mechanism [to apply 3rd party provided policy for flows associated with an application](#). The policy may contain e.g. coordination information.

UC #1 Immersive multi-modal Virtual Reality (VR) application (3 of 3)

- As the asynchrony between different modalities increases, users' sense of presence and realism will decrease.
- e.g., A study on delay threshold between visual effect after tactile is 15 ms; If more than 15 ms, user's perception of realism will be broken. Thus, E2E latency of one direction (uplink or downlink air interface) is set for 5 ms.

Synchronization thresholds for immersive multi-modality VR applications

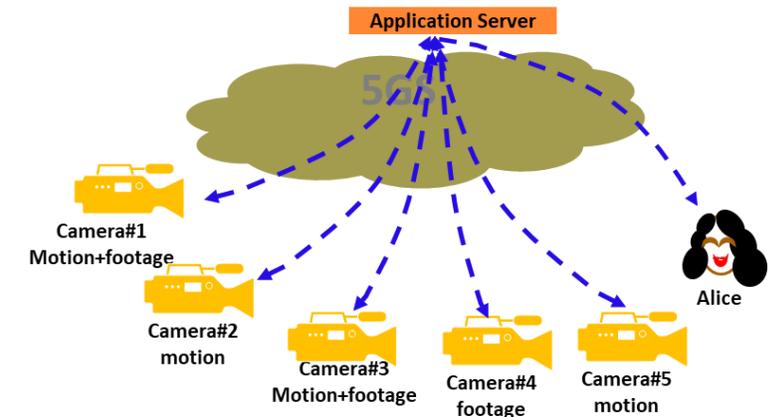
synchronization threshold (note 1)		
audio-tactile	audio delay: 50 ms	tactile delay: 25 ms
visual-tactile	visual delay: 15 ms	tactile delay: 50 ms

KPIs for tactile and multi-modal communication service

Use Cases	Characteristic parameter (KPI)			Influence quantity			Remarks	
	Max allowed end-to-end latency	Service bit rate: user-experienced data rate	Reliability	Message size (byte)	# of UEs	UE Speed		Service Area
Immersive multi-modal VR (UL: device → application sever)	5 ms (note 2)	16 kbit/s -2 Mbit/s (without haptic compression encoding); 0.8 - 200 kbit/s (with haptic compression encoding)	[99.9%] (without haptic compression encoding) [99.999%] (with haptic compression encoding)	1 DoF: 2-8 3 DoFs: 6-24 6 DoFs: 12-48 More DoFs can be supported by the haptic device	-	Stationary or Pedestrian	typically < 100 km ² (note 3)	Haptic feedback
	5 ms	< 1Mbit/s	[99.99%]	MTU	-	Stationary or Pedestrian	typically < 100 km ² (note 3)	Sensor information e.g. position and view information generated by the VR glasses
Immersive multi-modal VR (DL: application sever → device)	10 ms (note1)	1-100 Mbit/s	[99.9%]	1500	-	Stationary or Pedestrian	typically < 100 km ² (note 3)	Video
	10 ms	5-512 kbit/s	[99.9%]	50	-	Stationary or Pedestrian	typically < 100 km ² (note 3)	Audio
	5 ms (note 2)	16 kbit/s -2 Mbit/s (without haptic compression encoding); 0.8 - 200 kbit/s (with haptic compression encoding)	[99.9%] (without haptic compression encoding) [99.999%] (with haptic compression encoding)	1 DoF: 2-8 3 DoFs: 6-24 6 DoFs: 12-48	-	Stationary or Pedestrian	typically < 100 km ² (note 3)	Haptic feedback

UC #6 Live Event Selective Immersion

- AI Camera #1 to #5 are switched on and registered to the 5GS network to collect footage and predict motion of ball and player.
- The Application Server informs 5GS that UEs corresponding to Camera#1 to Camera#5 are subject to the service application, and provides QoS requirements of these UEs and the coordination policies for this multi-modal service for assistance from 5GS.
- Camera#1 is the primary camera placed in a location owning the best view. Camera#2 and #5 are for motion prediction.
- Network congestion happens from time to time. Based on the coordination policy of the multi-modal service, when the target QoS of Camera#1 can't be guaranteed, 5GS reduces QoS of Camera#4 to make sure QoS of Camera#1 is guaranteed; when the congestion is relieved, 5GS increases QoS of Camera#4 while target QoS of Camera#1 is still guaranteed.
- The network congestion gets more serious, and target QoS of Camera#2 can't be guaranteed. Since the motion data collected by Camera#2 is mandatory for motion prediction without which the motion prediction can't be made, 5GS releases resources of Camera#2 and Camera#5 based on the coordination policy of the multi-modal service.



TACMM: CR on TS 22.261 S1-214324

- Normative requirements and KPIs for tactile and multi-modal communication service is added to clause 6.43 of TS 22.261 (Service requirements for the 5G system)
- Requirement
 - The 5G system shall enable an authorized 3rd party to provide policy(ies) for flows associated with an application. The policy may contain e.g. the set of UEs and data flows, the expected QoS handling and associated triggering events, other coordination information.
 - The 5G system shall support a means to apply 3rd party provided policy(ies) for flows associated with an application. The policy may contain e.g. the set of UEs and data flows, the expected QoS handling and associated triggering events, other coordination information.

	synchronization threshold (note 1)	
audio-tactile	audio delay: 50 ms	tactile delay: 25 ms
visual-tactile	visual delay: 15 ms	tactile delay: 50 ms

Synchronization thresholds for immersive multi-modality VR applications

Use Cases	Characteristic parameter (KPI)			Influence quantity			Remarks	
	Max allowed end-to-end latency	Service bit rate: user-experienced data rate	Reliability	Message size (byte)	# of UEs	UE Speed		Service Area
Immersive multi-modal VR (UL: device → application sever)	5 ms (note 2)	16 kbit/s -2 Mbit/s (without haptic compression encoding); 0.8 - 200 kbit/s (with haptic compression encoding)	[99.9%] (without haptic compression encoding) [99.999%] (with haptic compression encoding)	1 DoF: 2-8 3 DoFs: 6-24 6 DoFs: 12-48 More DoFs can be supported by the haptic device	-	Stationary or Pedestrian	typically < 100 km ² (note 3)	Haptic feedback
	5 ms	< 1Mbit/s	[99.99%]	MTU	-	Stationary or Pedestrian	typically < 100 km ² (note 3)	Sensor information e.g. position and view information generated by the VR glasses
Immersive multi-modal VR (DL: application sever → device)	10 ms (note 1)	1-100 Mbit/s	[99.9%]	1500	-	Stationary or Pedestrian	typically < 100 km ² (note 3)	Video
	10 ms	5-512 kbit/s	[99.9%]	50	-	Stationary or Pedestrian	typically < 100 km ² (note 3)	Audio
	5 ms (note 2)	16 kbit/s -2 Mbit/s (without haptic compression encoding); 0.8 - 200 kbit/s (with haptic compression encoding)	[99.9%] (without haptic compression encoding) [99.999%] (with haptic compression encoding)	1 DoF: 2-8 3 DoFs: 6-24 6 DoFs: 12-48	-	Stationary or Pedestrian	typically < 100 km ² (note 3)	Haptic feedback

KPIs for tactile and multi-modal communication service

XR (Extended Reality) And Media Services

- **SA2 Study (FS_XRM: TR 23.XXX) '21.12→'22.09**
- Objectives
 - Enhancements for supporting **multi-modality service**:
 - Study whether and how to **enable delivery of related tactile and multi-modal data** (e.g., audio, video and haptic data related to a specific time) to the user at a similar time.
 - Enhancements of network exposure to support interaction between 5GS and application:
 - Study whether and how interaction between AF and 5GS is needed **for application synchronization and QoS policy coordination among multiple UEs or between multiple QoS flows per UE**.
 - Study **exposure of 5GS QoS information (e.g., QoS capabilities) and network conditions to the Application** to enable quick codec/rate adaptation help to provide desired QoE (e.g. such as assist in alleviating 5GS congestion).
 - NOTE: Parameters for exposure may coordinate with RAN and **SA4**.
 - Study whether and how the following QoS and policy enhancements for XR service and media service transmission are performed (traffic characteristics, different QoS handling for different importance of media unit, etc).
 - Study potential enhancements of Mobility and **power management considering traffic pattern of media services**:
 - Support handover enhancement to minimizing service disruption.
 - Power saving enhancement e.g. support trade-off of throughput/latency/reliability considering device battery life, whether and how to enhance CDRX, considering XR/media traffic pattern.
 - Whether and how to define **a network slice type supporting media services**.

Observation Points

- SA4 XR related works in Rel-18~ may have relationships with TACMM (SA1) and XRM (SA2).
- The UC #1 "Immersive multi-modal Virtual Reality (VR) application" and the UC #6 "Live event selective immersion" provides multi-sources with different QoS flows for one application. The use case could possibly be considered for a network slicing study in SA4.
- The UC #6 "Live event selective immersion" seems relevant to FS_NPN4AVProd.
- From an SA4 perspective, both TACMM and XRM may be relevant to delivery, synchronization and QoS of multi-modal media data.
- Mobility and power management in the XRM objectives may be relevant to successors of 5GSTAR such as MeCAR and Split_XR.