Radiocommunication Study Groups

Source: Documents 5D/TEMP/655 (Rev.1)

Document 5D/XXXX-E
2 Oct 2022
English only

IAFI

FURTHER UPDATES TO WORKING DOCUMENT TOWARD PRELIMINARY DRAFT NEW REPORT ON ITU-R [IMT MULTIMEDIA]

1 Introduction

During the 41st meeting of Working Party 5D, work was continued on a new deliverable to address new multimedia related capabilities if IMT-2020 and a new drafting group created during the 39th meeting of WP5D under SWG Specific Applications continued further development of the working document.

2 Proposal

Further changes are proposed to the working document as well as the workplan of this deliverable.

____

Annex 3.3 to Document 5D/1361-E

WORKING DOCUMENT TOWARD A PRELIMINARY DRAFT NEW REPORT ITU-R M. [IMT.MULTIMEDIA]

(Question ITU-R 262/5)

[Editor's note: As this document is developed, WP5D may consider incorporating any overlapping information into Report ITU-R M.2373, as appropriate.]

(20YY)

TABLE OF CONTENTS

[Editor's note: TBD]

1. Scope:

This Report addresses the capabilities of IMT-2020 to distribute multimedia content such as video, audio, text and graphics, including support for real-time multimedia interactive applications. This report also addresses the capabilities of IMT-2020 user devices and base stations to support such multimedia communications with low latency and wider transmission bandwidth.

This new report complements Report ITU-R M.2373 on "Audio-visual capabilities and applications supported by terrestrial IMT systems" which addresses the capabilities of IMT systems for delivering audio-visual services to the consumers and also covers some aspects of production of audio-visual content.

2 Introduction

[Editor's note: This section addresses overall information on applications for multimedia content as emerging applications by IMT-2020.]

Multimedia applications include network video, digital magazine, digital newspaper, digital radio, social media, mobile TV, digital TV, touch media, etc., that are enabled by IMT-2020 technologies. Beyond the traditional media service, the new media application not only supports accurate delivery of content, but also supports real-time interaction and real-time uploading of self media content. The users can be both consumers and producers of new media content.

These applications for multimedia content include but are not limited to:

– audio-visual applications,

– network video applications,

– digital online magazine applications,

– digital online newspaper applications,

– internet radio applications,

– social media applications,

– mobile internet TV applications,

– internet TV applications,

– touch media applications,

– on line information distribution applications,

– on-demand video applications,

– imaging and audio distribution applications,

– content dissemination applications,

– file delivery application

3 Relevant ITU-R Recommendations and Reports

[Editor's note: To be added]

4 Acronyms

[Editor's note: To be added]

5 Trends and market demands of applications for multimedia content supported by IMT-2020 technologies

[Editor's note: This section addresses global trends and demands for applications for multimedia content using IMT-2020 systems from external organizations as well as ITU-R members.]

IMT-2020 is expected to revolutionize the mobile experience with much faster, always-on, always connected, and responsive mobile Internet. It is noted that more than half of the mobile Internet traffic during last few years has been used for video downloads globally. This trend is expected to further increase in the future with IMT-2020. IMT-2030 is expected to connect the user to a high-definition virtual world on his mobile device. High-speed streaming of 4K videos only takes a few seconds and it can support crystal clear audio clarity. Live events can be streamed via a IMT-2020 radio network with high definition. HD TV channels can be accessed on mobile devices without any interruptions. The entertainment industry will hugely benefit from the IMT-2020 wireless networks.

IMT-2020 is expected to provide upto 120 frames per second, high resolution, and higher dynamic range video streaming without interruption. The audio-visual experience with IMT-2020 will be exceptionally good. Augmented reality and virtual reality require HD video with low latency. IMT-2020 network is powerful enough to power AR and VR with an amazing virtual experience. HD virtual reality games are getting popular, and many companies are investing in VR-based gaming. High speed IMT network can offer a better gaming experience with high-speed Internet.

5 Technical characteristics of IMT-2020 technologies for multimedia communication

[Editor's note: This section addresses technology characteristics of IMT-2020 technologies for multimedia communication from external organizations as well as ITU-R members.]

[5D/1065-CHN&KOR]

Multicast and broadcast services (MBS) in IMT-2020 has been one of the most important feature required by IMT system to tackle the proliferation of video services, ad-hoc multicast/broadcast streams, software delivery over wireless, group communications and broadcast/multicast IoT applications.

It aims to provide a flexible and dynamic allocation of radio resources between unicast and multicast services within the network as well as support for a stand-alone deployment of multicast/broadcast network, to offer multimedia services (including new media, audio, and visual applications) over specific geographic areas (e.g., spanning a limited number of base stations) as well as wide geographic are respectively.

The IMT-2020 specification support for Multicast and Broadcast services in 3GPP was developed step-by-step. The IMT-2020 technologies consisting of 3GPP 5G-SRIT and 5G RIT MBS (including LTE based 5G Terrestrial Broadcast and NR MBS) which allows the MT-2020 technologies to efficiently deliver Multicast and broadcast services.

The LTE based 5G terrestrial Broadcast finished in 3GPP Release 16 is based on legacy LTE eMBMS that can be traced back to 3GPP Release 8&9 which lay down the foundation for cellular broadcast. LTE based 5G terrestrial Broadcast is able to apply 100% of radio resources of one or more radio carriers for the delivery of broadcast/multicast content in a stand-alone multicast/broadcast network comprising of multiple cells with inter-site distances of up to 200 km.

The NR MBS in 3GPP Release 17 enables such a service over a specific geographic areas which further enables a more efficient and effective delivery system for real-time and streaming multicast/broadcast content. NR MBS is able to carry IP multicast as well as Ethernet multicast packets with better QoS support using dynamic delivery mode switching.

[5D/1334-China]:

Beyond the MBS in IMT-2020, from 3GPP Release 15, it has been developed related IMT-2020 technologies to offer very high uplink throughput, lower latency and/or higher capacity to support real-time multi-media information interaction and real-time uploading of self-media content.

Interference coordination is one of key characteristics. If non-synchronized frame structures are deployed for public and non-public networks, the interference between gNB and gNB or UE and UE would cause big impacts on the overall system performance. Interference coordination specified in 3GPP Release-16 could mitigate such interference.

Carrier aggregation (CA) and dual connectivity (DC) are the straightforward and effective ways to boost uplink throughput by directly increasing the uplink frequency resources and possibly cell capacity. With CA technology, a UE can receive or transmit on one or multiple contiguous/non-contiguous component carriers. The CA was firstly introduced in LTE Release 10. CA in IMT-2020 was specified from 3GPP Release 15 and support maximum 16 carriers with 400MHz each, thus it can support up to 6.4GHz bandwidth. It is continuously developed in 3GPP Release 16, Release17 and so on to support more flexible schedule and interference mitigation. The Multi-Radio Dual connectivity (MR-DC) was introduced in 3GPP Release 15. It is comprised of LTE+NR DC, NR+NR DC, NR+LTE DC etc. The RATs utilized in the first node and second node can be different. It is also continuously enhanced in 3GPP Release 15, Release 17 and so on to improve radio resource efficiency and reduce latency.

UL MIMO is another technology to improve up link data rate. It was specified for IMT-2020 from 3GPP Release 15. It introduces CSI feedback and reference signalling design which are more flexible than LTE. With high resolution codebook and beam forming characteristics, the data throughput and capacity of IMT-2020 system are expected to be met. In 3GPP Release 16, Release 17, it continuously developed CSI feedback, Coordinated Multiple Points (CoMP), beam management, power control, PAPR reference etc. to offer higher radio resource efficiency and communication performance for new media e.g. XR applications.

6 Use cases

[Editor's note: This section addresses potential use cases of applications for multimedia content *by IMT-2020 technologies.]*

6.1 Ultra-high-definition video/6.7 UHD Show

[5D/1334-China]

Live events (such as concerts, stage shows, sports events, etc.) require more details. For instance, in the concert, the 4K-based multi-screen and multi-camera live streaming can capture extremely vivid details such as subtle "micro-expression" changes, bringing a strong visual impact to the audience. Meanwhile, the application of HDR and wide color gamut technology will perfectly present the effect of live lighting and stage design to the audience. In band performances, multi-camera allows the audience to see their favorite players. In addition, ultra high-fidelity audio transmission is also necessary for stage shows. For large concerts (such as orchestras), multi-channel audio transmission with Hi-Fi music quality with inaudible noise and distortion, and a flat frequency response within the human hearing range, will allow listeners to experience a truly immersive experience.

With the support of UHD streaming by IMT-2020, the above scenarios will be feasible. All of the above features require huge bandwidth as a basis, and the IMT-2020 network can enable high-quality video and audio to be distributed to audiences with minimal delay.

6.2 Virtual reality (VR) panoramic video

6.3 Augmented reality (AR) images

6.4 Live Streaming

[5D/1334-China]:

Live streaming has become an important information distribution channel, which has gradually become a very popular form of entertainment. The popularity of live streaming is no less than the media that traditional entertainment platforms rely on (such as TV or Radio). For instance, top tier live streamers can attract millions of viewers when they are streaming. However, unlike those produced programs, there is no fixed schedule for live streaming, which means the live streamers could start their streaming anytime, anywhere. This characteristic will trigger an explosive surge in online viewing within a geographic area or within a period of time, which brings great challenges to the network transmission of the streaming platform. Therefore, the streaming platform has to pay high operating costs to expand bandwidth and IDC capacity to cope with the peak traffic demand. However, the massive number of equipment will cause a huge squandering of resources when the traffic drops. Moreover, the way of live streaming is evolving as well. For instance, the streaming video is evolving to ultra-high definition and immersive, which requires the higher video bit rate and stricter latency tolerance.

IMT-2020 network empowered streaming will be benefit to solve above issues. By integrating communication and information distribution application by supported IMT-2020, the super large access capacity can effectively address the concentration of viewing time and geography caused by super-influencer across the country or region, to solve the explosive traffic peak caused by influential contents. Within the ultra-high bandwidth and excellent latency performance of IMT-2020 networks, the realization difficulty of higher definition and real-time streaming will also be greatly reduced.

6.5 Live Commerce

[5D/1334-China]:

The term "Live commerce" refers to the use of live webcast technology to carry out new sales methods such as online display of merchandise, customer Q&A, and shopping guide sales through the Internet platform. Manufacturers usually use professional platforms to build their live streaming booth to sell a variety of products by streamers. Unlike traditional home shopping channel, Live commerce could interact well with customers. Furthermore, live commerce bypasses the traditional intermediate channels such as dealers so it could offer more favorable prices. It achieves seamless connectivity between merchandise and consumers, and attracts consumers' attention through the promotion of reliable quality and cost-effective products. This new way of shopping effectively reduces the cost of trust in the consumer choice process, and promotes purchase behavior.

Considering the characteristics of live commerce, the following interaction requirements should be met:

– The live streamers can receive the feedback from the viewer and response further with supplemental product introduction, which help enhance the viewer's confidence in the quality of the product.
– The live streamers can use the virtual reality (VR) / augmented reality (AR) to provide product trials and related functions, allowing customers to "experience" the product.
– The viewer can order and pay for the products online in real-time through the payment methods of live streamers (Payment link/QR code, etc.).
– The live streamers can conduct online prize draws or mini-games, to improve viewer stickiness.

Therefore, in a live commerce, distributing low-latency and high-quality live video to a large number of viewers, and the interaction with the viewer through the network will be a strong demand. Through the integration of empowered streaming, as well as the traditional wired networks will be a good practice to meet the above requirement.

6.6 Enhanced Venues for Live Events

[5D/1334-China]:

With the construction of IMT-2020 networks, the concept of "Smart Venue" that empower large venues with IMT-2020 networks and provide a new viewing experience to live events or activities has emerged. In top events, new viewing methods such as multi-view point and free-view point have been introduced, bringing a different experience to the website viewer and spectator. Taking "multi-view point" as an example, by deploying multiple cameras on stadium, content service providers can allow the viewer watch the game with different viewing angles from the regular streaming.

In addition to providing a diverse experience for viewers, another core concept of "Smart Venue" is to provide spectator with live services. For instance, during a game or event, the spectator can choose a different perspective from the current seat to watch the game via the client. During the intermission, the spectator could watch the replay of highlights or behind-the-scenes.

Generally, a large event often has tens or even hundreds of thousands of spectators, and millions of viewers. Such a dense population leads to huge construction and O&M costs for communication network coverage. When only using IMT-2020 networks to provide live services to users, service providers will face the great challenges of ultra low latency and ultra-dense connectivity.

5G empowered streaming could provide service to users in or out of the venue, which will effectively get rid of the network transmission bandwidth limitation and greatly reduce the cost of venue network construction. Meanwhile, the venue owners or event organizers will be able to leverage 5G empowered streaming to bring viewers (or spectators) an unprecedented experience.

6.7 Events Live streaming

[5D/1334-China]:

The production of traditional large sports events is mainly completed by sportscast production systems and equipment such as sportscast vans. However, ultra-high-definition sports video has gradually become the mainstream, and the traditional sportscast system is becoming more and more inadequate in terms of transmission bit rate and delay. At the same time, the content of live events is becoming more and more complex, with a large number of cameras, microphones and sensors working at the same time, and a large amount of data need to be transmitted back to the control center. In addition, as events become more globalized, sportscast resources need to be quickly deployed to all corners of the world in a very short period of time, and traditional methods (such as sportscast van) are difficult to achieve. In fact, the event broadcast production is gradually developing towards centralized, remote and lightweight. The IMT-2020 network empowered streaming has the characteristics of low latency, high bandwidth and wide coverage, which could improve the sportscast production process, optimize resource allocation, and have the ability to quickly distribute content to viewers.

In particular, the high bandwidth and low latency transmission of 5G enables the signals of some independent camera (such as Video Adjudication System) to be backhauled through the 5G network, which is conducive to the lightweight and rapid deployment of live streaming equipment. For cameras with special mobility (such as Spidercam), the use of 5G signal transmission will expand the moving range of the cameras in the venue. In addition, the use of high speed streaming to quickly distribute the camera signals from multi-viewing point, AR/VR signals, leaderboard and other content generated on the event.

Specifically, the following aspects could be enhanced:

– Short-distance transmission from multi-camera to live streaming system: multiple cameras (such as Steadicam, panoramic VR, AR hardware, etc.) are simultaneously transmitted to the sportscast van through IMT-2020for live streaming production. The Single-channel high-definition video (mainly 1920*1080i) is encoded with a bit rate of 15Mbps, and the uplink bit rate of the transmission channel will not be lower than this value. The relative delay between multiple cameras (same signal types) is less than 40ms to meet synchronization requirements.

– Lightweight deployment of live streaming equipment: IMT-2020 has extremely low delay for both DL&UL transmission, multiple signals (such as video/audio/data) can be transmitted to the remote production center through wireless transmission. Therefore, the transmission of control, tally, and voice signals can ensure extremely low delay in both uplink and downlink. These devices only need to be connected to the IMT-2020 network, whereas in the past they required thousands of cables to connect to each other, which avoids the complicated deployment process.

– Multi-content streaming services: Through high speed streaming, various content (such as multi-camera, VR, Box score, etc.) can be distributed to a large number of spectators in real time, allowing them to fully grasp the details of the event and provide a better viewing experience.

7 Current capabilities of applications for multimedia content supported by IMT-2020 technologies

[Editor's note: This section addresses technical and operational requirements for applications for multimedia content by IMT-2020 technologies, based on information from ITU-R members and/or external organizations.]

7.1 Necessary Requirement / Capabilities

[Editor's note: This section introduces required capabilities to support multimedia communication e.g. broadcast, multicast, higher throughput for UL, URLLC for XR like interaction etc.]

7.1.1 System capabilities requirements for broadcast/multicast

[5D/1065-CHN&KOR]

The following set of requirements complement the requirements listed in 3GPP TS 22.146, TS 22.246 and TS 22.101.

The IMT-2020 shall support operation of downlink only broadcast/multicast over a specific geographic area (e.g. a cell sector, a cell or a group of cells).

The IMT-2020 shall support operation of a downlink only broadcast/multicast system over a wide geographic area in a spectrally efficient manner for stationary and mobile UEs.

The IMT-2020 shall enable the operator to reserve 0% to 100% of radio resources of one or more radio carriers for the delivery of broadcast/multicast content.

The IMT-2020 shall allow the UE to receive content via a broadcast/multicast radio carrier while a concurrent data session is ongoing over another radio carrier.

The IMT-2020 shall be able to support broadcast/multicast of UHD streaming video (e.g. 4K/8K UHD).

NOTE 1: Taking into account the bandwidth needs for different streaming video resolution.

The IMT-2020 network shall allow the operator to configure and broadcast multiple quality levels (i.e. video resolutions) of broadcast/multicast content for the same user service in a stand-alone 3GPP based broadcast/multicast system.

The IMT-2020 network shall support parallel transfer of multiple quality levels (i.e. video resolutions) of broadcast/multicast content for the same user service to the same UE taking into account e.g. UE capability, radio characteristics, application information.

The IMT-2020 shall support parallel transfer of multiple multicast/broadcast user services to a UE.

The IMT-2020 shall support a stand-alone multicast/broadcast network comprising of multiple cells with inter-site distances of up to 200 km.

The IMT-2020 shall support multicast/broadcast via a 5G satellite access network, or via a combination of a 5G satellite access network and other 5G access networks.

The IMT-2020 shall be able to setup or modify a broadcast/multicast service area within [1s].

NOTE 2: For MCPTT related KPIs see 3GPP TS 22.179, clause 6.15.

The IMT-2020 shall be able to apply QoS, priority and pre-emption to a broadcast/multicast service area.

The IMT-2020 shall support downlink parallel transfer of the same content, via broadcast/multicast and/or unicast, such that all receiver group members in a given area receive the media at the same time according to user perception.

NOTE 3: In this context user perception refers to a difference in delay of typically less than 20 ms.

The IMT-2020 shall support a mechanism to inform a media source of relevant changes in conditions in the system (e.g. capacity, failures).

The IMT-2020 shall provide means for a media source to provide QoS requirement requests to the broadcast/multicast service.

The IMT-2020 shall provide means for the broadcast/multicast service to inform the media source of the available QoS, including modification of available QoS characteristics and availability of the broadcast/multicast service.

The IMT-2020 shall be able to support broadcast/multicast of voice, data and video group communication, allowing at least 800 concurrently operating groups per geographic area.

NOTE 4: In this context "concurrently operating groups" means that the associated media streams are delivered concurrently.

7.1.2 Radio access capabilities requirements for broadcast/multicast

[5D/1065-CHN&KOR]

More specifically, the necessary radio access technology (RAT) capabilities include:

The IMT-2020 shall support existing Multicast/Broadcast services (e.g. download, streaming, group communication, TV, etc.).

The IMT-2020 shall support dynamic adjustment of the Multicast/Broadcast area based on e.g. the user distribution or service requirements.

The IMT-2020 shall support concurrent delivery of both unicast and Multicast/Broadcast services to the users.

The IMT-2020 shall support efficient multiplexing with unicast transmissions in at least frequency domain and time domain.

The IMT-2020 shall support static and dynamic resource allocation between Multicast/Broadcast and unicast; the new RAT shall in particular allow support of up to 100% of DL resources for Multicast/Broadcast (100% meaning a dedicated MBMS carrier).

The IMT-2020 shall support Multicast/Broadcast network sharing between multiple participating MNOs, including the case of a dedicated MBMS network.

The IMT-2020 shall make it possible to cover large geographical areas up to the size of an entire country in SFN mode with network synchronization and shall allow cell radii of up to 100 km if required to facilitate that objective. It shall also support local, regional and national broadcast areas.

The IMT-2020 shall support Multicast/Broadcast services for fixed, portable and mobile UEs. Mobility up to 250 km/h shall be supported.

The IMT-2020 shall leverage usage of RAN equipment (hard- and software) including e.g. multi-antenna capabilities (e.g. MIMO) to improve Multicast/Broadcast capacity and reliability.

The IMT-2020 shall support Multicast/Broadcast services for mMTC devices.

NR system enables resource efficient delivery of multicast/broadcast services (MBS).

7.1.3 Requirements for real-time media information interaction and media content uploading

[5D/1334-China]:

Refer 3GPP TS22.261, the Audio-visual interaction is characterised by a human being interacting with the environment or people, or controlling a UE, and relying on audio-visual feedback which is also the characteristics of new media interaction and uploading. To support VR environments with low motion-to-photon capabilities, the IMT-2020 shall support:

– motion-to-photon latency in the range of 7 ms to 15ms while maintaining the required resolution of up to 8k giving user data rate of up to [1Gbit/s] and

– motion-to-sound delay of [< 20 ms].

NOTE: The motion-to-photon latency is defined as the latency between the physical movement of a user's head and the updated picture in the VR headset. The motion-to-sound latency is the latency between the physical movement of a user's head and updated sound waves from a head mounted speaker reaching their ears.

To support interactive task completion during voice conversation, the IMT-2020 shall support low-delay speech coding for interactive conversational services (100 ms, one-way mouth-to-ear).

Due to the separate handling of the audio and video component, the IMT-2020 will have to cater for the VR audio-video synchronisation in order to avoid having a negative impact on the user experience (i.e. viewers detecting lack of synchronization). To support VR environments the IMT-2020 shall support audio-video synchronisation thresholds:

– in the range of [125 ms to 5 ms] for audio delayed and

– in the range of [45 ms to 5 ms] for audio advanced.

When it comes to implementation of applications containing AR/VR components, the requirements on the IMT-2020 network could depend on architectural choices implementing these services. Note 3 in 3GPP TS22.261 table 7.1-1 gives an example on such dependences for a VR application in IMT-2020. Table 7.6.1-1 illustrates additional AR/VR use cases and provides more corresponding requirements on the IMT-2020.

– Cloud/Edge/Split Rendering – Cloud/Edge/Split Rendering is characterised by the transition and exchange of the rendering data between the rendering server and device.

– Gaming or Training Data Exchanging – This use case is characterised by the exchange of the gaming or training service data between two IMT-2020 networks connected AR/VR devices.

– Consume VR content via tethered VR headset – This use case involves a tethered VR headset receiving VR content via a connected UE; this approach alleviates some of the computation complexity required at the VR headset, by allowing some or all decoding functionality to run locally at the connected UE. The requirements in the table below refer to the direct wireless link between the tethered VR headset and the corresponding connected UE.

TS22.261 TABLE 7.1-1

Performance requirements for high data rate and traffic density scenarios.

	Scenario	**Experienced data rate (DL)**	**Experienced data rate (UL)**	**Area traffic capacity (DL)**	**Area traffic capacity (UL)**	**Overall user density**	**Activity factor**	**UE speed**	**Coverage**
3	Indoor hotspot	1 Gbit/s	500 Mbit/s	15 Tbit/s/km^2	2 Tbit/s/km^2	250 000/km^2	Note 2	Pedestrians	Office and residential (Note 2) (Note 3)

NOTE 1: For users in vehicles, the UE can be connected to the network directly, or via an on-board moving base station.

NOTE 2: A certain traffic mix is assumed; only some users use services that require the highest data rates [2].

NOTE 3: For interactive audio and video services, for example, virtual meetings, the required two-way end-to-end latency (UL and DL) is 2-4 ms while the corresponding experienced data rate needs to be up to 8K 3D video [300 Mbit/s] in uplink and downlink.

NOTE 4: These values are derived based on overall user density. Detailed information can be found in [10].

NOTE 5: All the values in this table are targeted values and not strict requirements.

3GPP TS22.261 TABLE 7.6.1-1

KPI Table for AR/VR high data rate and low latency service

Use Cases	Characteristic parameter (KPI)			Influence quantity		
	Max allowed end-to-end latency	**Service bit rate: user-experienced data rate**	**Reliability**	**# of UEs**	**UE Speed**	**Service Area (Note 2)**
Cloud/Edge/Split Rendering (Note 1)	5 ms (i.e. UL+DL between UE and the interface to data network) (Note 4)	0,1 to [1] Gbit/s supporting visual content (e.g. VR based or high definition video) with 4K, 8K resolution and up to120 frames per second content.	99,99 % in uplink and 99,9 % in downlink (Note 4)	-	Stationary or Pedestrian	Countrywide
Gaming or Interactive Data Exchanging (Note 3)	10ms (Note 4)	0,1 to [1] Gbit/s supporting visual content (e.g. VR based or high definition video) with 4K, 8K resolution and up to120 frames per second content.	99,99 % (Note 4)	≤ [10]	Stationary or Pedestrian	20 m x 10 m; in one vehicle (up to 120 km/h) and in one train (up to 500 km/h)
Consumption of VR content via tethered VR headset (Note 6)	[5 to 10] ms (Note 5)	0,1 to [10] Gbit/s (Note 5)	[99,99 %]	-	Stationary or Pedestrian	–

NOTE 1: Unless otherwise specified, all communication via wireless link is between UEs and network node (UE to network node and/or network node to UE) rather than direct wireless links (UE to UE).

NOTE 2: Length x width (x height).

NOTE 3: Communication includes direct wireless links (UE to UE).

NOTE 4: Latency and reliability KPIs can vary based on specific use case/architecture, e.g. for cloud/edge/split rendering, and can be represented by a range of values.

NOTE 5: The decoding capability in the VR headset and the encoding/decoding complexity/time of the stream will set the required bit rate and latency over the direct wireless link between the tethered VR headset and its connected UE, bit rate from 100 Mbit/s to [10] Gbit/s and latency from 5 ms to 10 ms.

NOTE 6: The performance requirement is valid for the direct wireless link between the tethered VR headset and its connected UE.

3GPP TS22.263 specifies performance requirements for Video and audio production applications in Table 6.2.1-1

3GPP TS22.263 TABLE 6.2.1-1

Performance requirements of professional low-latency periodic deterministic audio transport service

Profile	**# of active UEs**	**UE Speed**	**Service Area**	**E2E latency (Note 1)**	**Transfer interval (Note 1)**	**Packet error rate (Note 2, Note 3)**	**Data rate UL**	**Data rate DL**
Music Festival	200	10 km/h	500 m x 500 m	750 µs	250 µs	10^{-6}	500 kbit/s	-
	100	10 km/h	500 m x 500 m	750 µs	250 µs	10^{-6}	-	1 Mbit/s
Musical	30	50 km/h	50 m x 50 m	750 µs	250 µs	10^{-6}	500 kbit/s	-

	20	50 km/h	50 m x 50 m	750 µs	250 µs	10^{-6}	-	1 Mbit/s
	10	-	50 m x 50 m	750 µs	250 µs	10^{-6}	-	500 kbit/s
Semi-professional	10	5 km/h	5 m x 5 m	750 µs	250 µs	10^{-6}	100 kbit/s	-
	10	5 km/h	5 m x 5 m	750 µs	250 µs	10^{-6}	-	200 kbit/s
	2	-	5 m x 5 m	750 µs	250 µs	10^{-6}	-	100 kbit/s
AV production	20	5 km/h	30 m x 30 m	750 µs	250 µs	10^{-6}	1.5 Mbit/s	-
	10	5 km/h	30 m x 30 m	750 µs	250 µs	10^{-6}	-	3 Mbit/s
Audio Studio	30	-	10 m x 10 m	750 µs	250 µs	10^{-6}	5 Mbit/s	-
	10	5 km/h	10 m x 10 m	750 µs	250 µs	10^{-6}	-	1 Mbit/s

NOTE 1: Transfer interval refers to periodicity of the packet transfers. It has to be constant during the whole operation. The value given in the table is a typical one, however other transfer intervals are possible as long as the end-to-end latency is ≤ (1ms – Transfer interval).

NOTE 2: Packet error rate is related to a packet size of (transfer interval × data rate). Packets that do not conform with the end-to-end latency are also accounted as error.

NOTE 3: The given requirement for a packet error rate assumes a uniform error distribution. The requirement for packet error rate is stricter if packet errors occur in bursts.

3GPP TS22.263 TABLE 6.2.1-3

Performance requirements for low latency video.

Profile	# of active UEs	UE Speed	Service Area	E2E latency	Packet error rate (Note 1)	Data rate UL	Data rate DL
Uncompressed UHD video	1	0 km/h	1 km^2	400 ms	10^{-10} UL 10^{-7} DL	12 Gbit/s	20 Mbit/s
Uncompressed HD video	1	0 km/h	1 km^2	400 ms	10^{-9} UL 10^{-7} DL	3 .2 Gbit/s	20 Mbit/s
Mezzanine compression UHD video	5	0 km/h	1000 m^2	1 s	10^{-9} UL 10^{-7} DL	3 Gbit/s	20 Mbit/s
Mezzanine compression HD video	5	0 km/h	1000 m^2	1 s	10^{-9} UL 10^{-7} DL	1 Gbit/s	20 Mbit/s
Tier one events UHD	5	0 km/h	1000 m^2	1 s	10^{-9} UL 10^{-7} DL	500 Mbit/s	20 Mbit/s
Tier one events HD	5	0 km/h	1000 m^2	1 s	10^{-8} UL 10^{-7} DL	200 Mbit/s	20 Mbit/s
Tier two events UHD	5	7 km/h	1000 m^2	1 s	10^{-8} UL 10^{-7} DL	100 Mbit/s	20 Mbit/s

Profile	# of active UEs	UE Speed	Service Area	E2E latency	Packet error rate (Note 1)	Data rate UL	Data rate DL
Tier two events HD	5	7 km/h	1000 m^2	1 s	10^{-8} UL 10^{-7} DL	80 Mbit/s	20 Mbit/s
Tier three events UHD (Note 2)	5	200 km/h	1000 m^2	1 s	10^{-7} UL 10^{-7} DL	20 Mbit/s	10 Mbit/s
Tier three events HD (Note 2)	5	200 km/h	1000 m^2	1 s	10^{-7} UL 10^{-7} DL	10 Mbit/s	10 Mbit/s
Remote OB	5	7 km/h	1000 m^2	6 ms	10^{-8} UL 10^{-7} DL	200 Mbit/s	20 Mbit/s

NOTE 1: Packets that do not conform with the end-to-end latency are also accounted as error. The packet error rate requirement is calculated considering 1500 B packets, and 1 packet error per hour is $10^{-5}/(3*x)$, where x is the data rate in Mbps.

NOTE 2: Could use either professional equipment or mobile phone equipped with dedicated newsgathering app

3GPP TS22.263 TABLE 6.2.1-4

Performance requirements for airborne base stations for NPN.

Profile	# of active UEs	UE Speed	Service Area	E2E latency	Packet error rate (Note 1)	Data rate UL	Data rate DL
NPN ground to air UHD up Link	10	500 km/h	700 km2 x 6000 m (Note 2)	40 ms	10^{-8} UL 10^{-7} DL	100 Mbit/s	20 Mbit/s
NPN ground to air HD up link	10	500 km/h	700 km2 x 6000 m (Note 2)	40 ms	10^{-8} UL 10^{-7} DL	80 Mbit/s	20 Mbit/s
NPN air to ground UHD down Link	2	500 km/h	700 km2 x 6000 m (Note 2)	40 ms	10^{-7} UL 10^{-8} DL	20 Mbit/s	100 Mbit/s
NPN air to ground HD down link	2	500 km/h	700 km2 x 6000 m (Note 2)	40 ms	10^{-7} UL 10^{-8} DL	20 Mbit/s	80 Mbit/s
NPN radio Camera UHD	10	200 km/h	1 km^2	3 ms	10^{-8} UL 10^{-7} DL	100 Mbit/s	20 Mbit/s
NPN radio camera HD	10	200 km/h	1 km^2	3 ms	10^{-8} UL 10^{-7} DL	80 Mbit/s	20 Mbit/s

NOTE 1: Packets that do not conform with the end-to-end latency are also accounted as error. The packet error rate requirement is calculated considering 1500 B packets, and 1 packet error per hour is $10^{-5}/(3*x)$, where x is the data rate in Mbps.

NOTE 2: 6000 m = height but in a cone formation (i.e. ground coverage with a circle of diameter 30 KM)

Further, with development of immersive multiple modal integrated with AR/VR, 3GPP also specified capability requirements for the IMT-2020 which will greatly enrich and enhance user experience of the immersive multimedia service.

3GPP TS22.261 TABLE 7.11-1

Multi-modal communication service performance requirements

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)**	**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) [40]	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 5)	Haptic feedback
	5 ms	< 1Mbit/s	99.99% [40]	1500	Stationary or Pedestrian	typically < 100 km^2 (note 5)	Sensing 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% [40]	1500	Stationary or Pedestrian	typically < 100 km^2 (note 5)	Video
	10 ms	5-512 kbit/s	99.9% [40]	50	Stationary or Pedestrian	typically < 100 km^2 (note 5)	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) [40]	1 DoF: 2-8 3 DoFs: 6-24 6 DoFs: 12-48	Stationary or Pedestrian	typically < 100 km^2 (note 5)	Haptic feedback

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)**	**UE Speed**	**Service Area**	
Immersive multi-modal navigation applications Remote Site → Local Site (DL)	50 ms [39]	16 kbit/s -2 Mbit/s (without haptic compression encoding); 0.8 - 200 kbit/s (with haptic compression encoding)	99.999 % [40]	1 DoF: 2 to 8 10 DoF: 20 to 80 100 DoF: 200 to 800	Stationary or Pedestrian	≤ 100 km^2 (note 5)	Haptic feedback
	<400 ms [39]	1- 100 Mbit/s	99.999 % [40]	1500	Stationary/ or Pedestrian,	≤ 100 km^2 (note 5)	Video
	<150 ms [39]	5-512 kbit/s	99.9 % [40]	50	Stationary or Pedestrian	≤ 100 km^2 (note 5)	Audio
	<300 ms	600 Mbit/s	99.9 % [40]	1500	Stationary or Pedestrian	≤ 100 km^2 (note 5)	VR
Immersive multi-modal navigation applications Local Site → Remote Site (UL)	<300 ms	12 kbit/s [26]	99.999 % [40]	1500	Stationary or Pedestrian	≤ 100 km^2 (note 5)	Biometric / Affective
	<400 ms [39]	1- 100 Mbit/s	99.999 % [40]	1500	Workers: Stationary/ or Pedestrian, UAV: [30-300mph]	≤ 100 km^2 (note 5)	Video
	<150 ms [39]	5-512 kbit/s	99.9 % [40]	50	Stationary or Pedestrian	≤ 100 km^2 (note 5)	Audio
	<300 ms	600 Mbit/s	99.9 % [40]	1500	Stationary or Pedestrian	≤ 100 km^2 (note 5)	VR

NOTE 1: Motion-to-photon delay (the time difference between the user's motion and corresponding change of the video image on display) is less than 20 ms, and the communication latency for transferring the packets of one audio-visual media is less than 10 ms, e.g. the packets corresponding to one video/audio frame are transferred to the devices within 10 ms.

NOTE 2: According to IEEE 1918.1 [40] as for haptic feedback, the latency is less than 25 ms for accurately completing haptic operations. As rendering and hardware introduce some delay, the communication delay for haptic modality can be reasonably less than 5 ms, i.e. the packets related to one haptic feedback are transferred to the devices within 10 ms.

NOTE 3: Haptic feedback is typically haptic signal, such as force level, torque level, vibration and texture.

NOTE 4: The latency requirements are expected to be satisfied even when multimodal communication for skillset sharing is via indirect network connection (i.e., relayed by one UE to network relay).

NOTE 5: 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.

Beyond these performance requirements for new media interaction, following service capabilities are required to support new media real-time interaction and media content uploading which are specified in 3GPP TS22.263.

The IMT-2020 enables an NPN for video, imaging and audio for professional applications.

The IMT-2020 network shall be able to provide a time reference information to a 3rd party application acting as a master clock with an accuracy of 1 microsecond.

The IMT-2020 shall be able securely reconnect within a short period of time (<1s) from UE starting first network reconnection attempt after the UE has detected a UE network connection loss.

The IMT-2020 shall support uplink and downlink service continuity maintaining acceptable performance requirements while switching between co-located PLMN and NPN (e.g., due to mobility).

The IMT-2020 shall support service continuity maintaining acceptable performance requirements: for an uplink stream while performing traffic steering, switching, and splitting among co-located PLMN(s) and NPN(s); for downlink while switching between co-located PLMN and NPN.

7.2 IMT-2020 capabilities

[5D/1065-CHN&KOR]

The necessary capabilities listed in above session are met by different IMT capabilities in different Releases and in different network components or systems (radio network, architectures enhancements, and application layer support.

7.2.1 Radio Access Network

[5D/1065-CHN&KOR]

7.2.1.1 Terrestrial Broadcast

3GPP technology defined in Rel-16 aims to deliver audio-visual services (including free-to-air services) in single frequency network (SFN) on stand-alone infrastructure, usually in High-Power High-Tower Single Frequency Networks to support larger inter-site distance (e.g., to allow cell radii of up to 100 km).

To allow cell radii of up to 100 km, numerology enhancement with a new 0.37 kHz subcarrier spacing and CP duration ~300 µs was introduced to support terrestrial broadcast with conventional broadcasting in medium power medium tower (MPMT) & HPHT. To allow broadcast service reception with UE mobility up to 250 km/h, numerology enhancement with a new 2.5 kHz wider sub-carrier spacing, CP duration ~100 µs with better Doppler resiliency. There are other related enhancement to support broader coverage for the control information like dedicated reference signals (RS) accompany each numerology, enhanced subframe structure, control channel, and less dense RS pattern with reducing overheads.

7.2.1.2 Flexible and dynamic resources allocation

3GPP Rel-17 aims to support flexible and dynamic resources allocation to enable both Multicast and Broadcast support in use cases of public safety and mission critical, V2X applications, transparent IPv4/IPv6 multicast delivery, IPTV, software delivery over wireless, group communications and IoT applications, instead of broader area terrestrial broadcast like services.

RAN basic functions for broadcast/multicast for UEs in RRC_CONNECTED state are to be supported with better reliability and better service continuity.

For multicast communication service, the same service and the same specific content data are provided simultaneously to a dedicated set of UEs (i.e., not all UEs in the Multicast service area are authorized to receive the data). A multicast communication service is delivered to the UEs using a multicast session.

A UE can receive a multicast communication service in RRC_CONNECTED state with mechanisms such as point to point (PTP) and/or point to multi-point (PTM) delivery, with a balance between network efficiency. HARQ feedback/retransmission can be applied to both PTP and PTM transmissions.

For PTM transmission there can be two HARQ feedback schemes: UE specific feedback and NACK only feedback, depending on network decision based on level of reliability requirement or network resources. Network is able to dynamically change Multicast service delivery between PTM and PTP with service continuity for a given UE.

Basic mobility with service continuity. Unicast like mobility mechanism is designed to offer the basic mobility, e.g., the Multicast session resources are to be established along with UE's mobility in the target RAN node which supports the MBS service.

For broadcast communication service, the same service and the same specific content data are provided simultaneously to all UEs in a geographical area (i.e., all UEs in the Broadcast service area are authorized to receive the data). A broadcast communication service is delivered to the UEs using a broadcast session. A UE can receive a broadcast communication service in RRC_IDLE, RRC_INACTIVE and RRC_CONNECTED state.

There are other enhancement including, UE are able to receive MBS with simultaneous operation in unicast reception; enhancement on RAN network interface like Xn, F1, and E1 interfaces; support on dynamic broadcast area.

[5D/1334-China]:

7.2.1.3 UL enhancement

A) UL MIMO

While UL MIMO offers the capability for reduction in overhead and/or latency, high-speed vehicular scenarios (e.g. a UE traveling at high speed on highways) at FR2 require more aggressive reduction in latency and overhead – not only for intra-cell, but also for L1/L2 centric inter-cell mobility. This also includes reducing the occurrence of beam failure events. Besides of it, enhancements for enabling panel-specific UL beam selection was investigated and specified continuously. This offers some potential for increasing UL coverage including. Then, channels other than PDSCH can benefit from multi-TRP transmission (as well as multi-panel reception) which also includes multi-TRP for inter-cell operations. This includes some new use cases such as UL dense deployment within a macro-cell and/or heterogeneous-network-type deployment scenarios. And, the SRS can be further enhanced for capacity and coverage.

B) UL Carrier Aggregation

Carrier aggregation was developed the unaligned frame boundary capability to provide more flexible beginning TX frame structure configurations among different carriers to offer higher uplink throughput and lower latency. Tx switching is to specify the dynamic switching mechanisms among two uplink carriers. Unaligned frame boundary and Tx switching can be implemented together to achieve larger uplink throughput for TDD CA operation. In one cell group, it supports PUCCH carrier switching semi-statically or dynamically. This could reduce the latency of PUCCH transmissions significantly for CA operation.

C) UL Dual connectivity (DC)

Dual connectivity is capable of Uplink power control i.e. to limit UE's transmission power to assure edge user's communication QoS, migration interference among the users, and to reduce energy consumption of the UE. Uplink DC is also capable of earlier measurement and fast recovery from MCG failure to reduce latency. Beyond these, it is capable of to indicate UE entering the third state to maintain context of the UE which will reduce configuration overhead.

4 Interference coordination

Interference coordination includes two aspects, i.e., cross-link interference (CLI) and remote interference management (RIM). For CLI, it is left to network implementation for gNB-gNB interference. The UE-UE interference coordination is capable of such as SRS-RSRP/CLI-RSSI based layer-3 CLI measurement and reporting, and network coordination mechanism for CLI with inter-gNB exchange of intended UL/DL configuration. RIM targets to migrate the interference occurring in specific weather conditions with the distance between aggressor gNB and victim gNB hundreds of kilometres. The RIM reference signal (RS) based on PRACH preamble-like RS is introduced for better interference measurement while the detailed remote interference mitigation mechanisms are left to implementation.

7.2.2 Architecture Enhancements

[5D/1065-CHN&KOR]

The MBS architecture follows the IMT-2020 architectural principles as defined in 3GPP TS 23.501, enabling distribution of the MBS data from the 5GS ingress to NG-RAN node(s) and then to the UE. The MBS architecture provides:

– Efficient usage of RAN and CN resources, with an emphasis on radio interface efficiency;
– Efficient transport for a variety of multicast and broadcast services.

MBS traffic is delivered from a single data source (e.g. Application Service Provider) to multiple UEs. Depending on many factors, there are several delivery methods which may be used to deliver the MBS traffic in the 5GS.

The 5G MBS also provides functionalities such as local MBS service, authorization of multicast MBS and QoS differentiation.

Between 5GC and NG-RAN, there are two possible delivery methods to transmit the MBS data:

– 5GC Individual MBS traffic delivery method: This method is only applied for multicast MBS session. 5GC receives a single copy of MBS data packets and delivers separate copies of those MBS data packets to individual UEs via per-UE PDU sessions, hence for each such UE one PDU session is required to be associated with a multicast session.
– 5GC Shared MBS traffic delivery method: This method is applied for both broadcast and multicast MBS session. 5GC receives a single copy of MBS data packets and delivers a single copy of those MBS packets to an NG-RAN node, which then delivers the packets to one or multiple UEs.

The 5GC Shared MBS traffic delivery method is required in all 5G MBS deployments. The 5GC Individual MBS traffic delivery method is required to enable mobility when there is an NG-RAN deployment with non-homogeneous support of 5G MBS.

For the multicast session, a single copy of MBS data packets received by the CN may be delivered via 5GC Individual MBS traffic delivery method for some UE(s) and via 5GC Shared MBS traffic delivery method for other UEs.

New network functionality are introduced to support MBS in 5GS:

– MB-SMF, Supporting MBS session management (including QoS control).

– MB-UPF, user plane function for MBS with Packet filtering of incoming downlink packets for multicast and broadcast flows, and QoS enforcement.

Other network functions are enhanced, e.g., SMF to support UE join/leave operation of Multicast session, AMF to support group paging, and signaling routing, PCF to provide policy information and QoS handling for MBS session.

[5D/1334-China]:

The integration of XR applications within the IMT-2020 System is approached following the model of 3GPP 5G Media Streaming as defined in 3GPP TS 26.501 [x]. Assume a 3GPP 5G-XR Application Provider being an XR Application provider that makes use of 5G System functionalities for its services. For this purpose, it provides a 5G-XR Aware Application on the UE to make use of a 3GPP 5G-XR client and network functions using network interfaces.

The QoS model is described in clause 5.7 of 3GPP TS 23.501 [x]. The 5G QoS model supports both:

– QoS Flows that require guaranteed flow bit rate (GBR QoS Flows)

– and QoS Flows that do not require guaranteed flow bit rate (Non-GBR QoS Flows).

The QoS model also supports Reflective QoS (see clause 5.7.5 of 3GPP TS 23.501 [8]).

A QoS Flow ID (QFI) is used to identify a QoS Flow in the IMT-2020. User Plane traffic assigned to the same QoS Flow within a Protocol Data Unit (PDU) Session receives the same traffic forwarding treatment (e.g. scheduling, admission threshold).

The QFI may be dynamically assigned or may be equal to the 3GPP 5G QoS Identifier (5QI). A QoS Flow may either be 'GBR', 'Non-GBR' or “Delay Tolerant GBR” depending on its QoS profile and it contains QoS parameters as follows:

– For each QoS Flow, the QoS profile includes the QoS parameters:
  - 3GPP 5G QoS Identifier (5QI); and
  - Allocation and Retention Priority (ARP).

– For each Non-GBR QoS Flow only, the QoS profile can also include the QoS parameter:
  - Reflective QoS Attribute (RQA).

– For each GBR QoS Flow only, the QoS profile also include the QoS parameters:
  - Guaranteed Flow Bit Rate (GFBR) - uplink (UL) and downlink (DL); and
  - Maximum Flow Bit Rate (MFBR) - UL and DL; and

– In the case of a GBR QoS Flow only, the QoS profile can also include one or more of the QoS parameters:
  - Notification control;
  - Maximum Packet Loss Rate - UL and DL

In Release 18, PDU Set concept is introduced in FS_XRM study to optimize the delivery of XRM service in 5GS. A PDU Set is composed of one or more PDUs carrying the payload of one unit of information generated at the application level (e.g., a frame or video slice for XRM Services), which are of same importance at application layer. All PDUs in a PDU Set are needed by the application layer to use the corresponding unit of information. In some cases, the application layer can still recover parts of the information unit, when some PDUs are missing.

The following key issues are currently under investigating:

WT#1: 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) with an application to the user at a similar time, focusing on the need for policy control enhancements (e.g. QoS policy coordination).

WT#2: Enhancements of network exposure to support interaction between 5GS and application：

– WT#2.1: 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.
– WT#2.2: 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).

WT#3: Study whether and how the following QoS and policy enhancements for XR service and media service transmission are performed:

– WT#3.1: Study the traffic characteristics of media service enabling improved network resources usage and QoE.
– WT#3.2: Enhance QoS framework to support media units granularity (e.g., video/audio frame/tile, Application Data Unit, control information), where media units consist of PDUs that have the same QoS requirements.
– WT#3.3: Support differentiated QoS handling considering different importance of media units. e.g., eligible drop packets belong to less important media units to reduce the resource wasting.
– WT#3.4: Whether and how to support uplink-downlink transmission coordination to meet RTT (Round-Trip Time) latency requirements between UE and N6 termination point at the UPF.
– WT#3.5: Potential policy enhancements to minimize the jitter, focusing on i.e. requirement provisioning from AF, extension of PCC rule.

WT#4: Study potential enhancements of Mobility and power management considering traffic pattern of media services:

– WT 4.1: void
– WT 4.2: 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.

8 Case study

[Editor's note: This section provides case studies from various countries associated with applications for multimedia content *in various usage scenarios supported by eMBB, mMTC and URLLC of terrestrial IMT-2020 systems]*

9 Summary

[Editor's note: To be described.]

____