IBC2024

3GPP SA4 -
THE MEDIA AND XR POWERHOUSE FOR 5G AND 6G

Frédéric Gabin[1], Jaeyeon Song[2], Gilles Teniou[3], Thomas Stockhammer[4]

[1]Dolby Labs, Paris, France
[2]Samsung, Seoul, South Korea
[3]Tencent, Rennes, France
[4]Qualcomm Germany GmbH, Munich, Germany

ABSTRACT

5G-based Media Distribution, developed by 3GPP, is getting ubiquitous, global and across different verticals. 3GPP SA4, the 3GPP working group for media technologies, is the powerhouse for developing the specifications relevant for content delivery, compression, quality-of-experience, codecs and formats. With the entrance into 5G advanced as of Release-18, 3GPP has addressed the needs for traditional content distribution, for emerging XR services, the integration of AI/ML into media workflows and the support for verticals for their media needs. This paper addresses a summary of completed Rel-18 technologies for media and XR related technologies and provides an overview on what is happening and expected in Rel-19, and possibly even beyond in the 6G era. For XR, normative specifications have been developed to address media capabilities for AR glasses including XR runtime, rendering capabilities, new formats, and codecs for audiovisual experiences. The integration into 5G delivery is of utmost importance to support functionalities including split rendering and distributed compute. Those technologies require smart distribution of compute resources across handheld devices, glasses or other HMDs (Head Mounted Displays), and network elements. Requirements and technologies such as low latency encoding and delivery, support for QoS (Quality of Service) and appropriate APIs (Application Programming Interfaces) open a platform for new Metaverse experiences. 3GPP SA4 also has dealt with upgrading functionalities for conversational experiences, including a new immersive voice and audio services (IVAS) codec, integration of webRTC into 3GPP, as well as using Avatars to augment traditional IMS-based calls. This paper provides a detailed overview of 3GPP (3rd Generation Partnership Project) Release-18 functionalities as well as an overview on reference implementations and prototypes that are developed in collaboration with 5G-MAG. Finally, the paper will provide an outlook to expected work in the next Releases, with a media vision into 6G.

INTRODUCTION

5G and its latest evolution called 5G-Advanced are seen as real game changers for the media industry, with the promises to provide faster, more reliable, and more responsive connectivity. This will enable the delivery of content more quickly and efficiently, which is particularly important in today's fast-paced media landscape. Higher bandwidth, ultralow latency and more reliable network configurations are at the core of the evolutions for which the benefits are two-fold: the optimization of the QoE for audiovisual services and the introduction of disruptive experiences such as eXtended Reality (XR).

This paper first introduces the latest developments in 3GPP with respect to media and XR, in particular 3GPP SA4, the 3GPP working group for media technologies, is the powerhouse for developing the specifications relevant for content delivery, compression, quality-of-experience, codecs and formats.

5G MEDIA: USE CASES, MARKET NEEDS, CHALLENGES AND OPPORTUNITIES

Approach

3GPP SA4 work is driven by use cases, requirements, challenges and opportunities that arise from continuous work with global service and technology providers. 3GPP SA4 work is also driven by input from other 3GPP working groups that rely on smooth integration of media services into 5G System architecture and radio services. Before diving into technology areas, a few words about the methodology and processes. 3GPP is organized in Releases, typically spanning over a period of 18 months. In June 2024, Rel-18 was completed, and Rel-19 was kicked off. Within the timeframe of a Release, several work topics are identified, and concrete objectives are agreed upon that are then mapped into updates to existing specifications or lead to the creation of new specifications. In many cases, the normative work phase is proceeded by studies that help to identify the exact needs, discuss and evaluate potentially different solutions, and provide a pre-determination of the exact scope of normative work. In certain cases, these studies are simple and can be done within the first few months of a Release. In other cases, the studies are extensive and may span the time frame of a Release.

3GPP’s work, and in particular SA4’s work is guided by diligent evaluation of market needs that may impact deployments within the next 1-3 years when the normative work is started. It is particularly guided by the feasibility to integrate the technologies into upcoming UEs and network infrastructures and supports interoperable and globally deployable technologies, addressing the need for economy of scale. As an example, the support for new codecs or profiles of new codecs results in significant changes in a hardware platform on UEs. Such integration for example requires understanding of required complexity, area size, power consumption, benefits and market needs. Beyond this, to add new technologies, it needs to be ensured that a detailed set of testing regimes as well as interoperability and conformance programs exist. Generally, 3GPP SA4 addresses the integration of new technologies into 3GPP specifications together with the evaluation of all the above issues above: market needs, maturity, benefits, interoperability and implement ability.

Addressing Third-Party Services

While 3GPP for a long time primarily served mobile network operator services, with the expansion to verticals as well as by the de-composition of services into “apps”, the support for third-party service providers by 3GPP specifications has become as relevant as operator services. To address the needs for third-party service support, 3GPP SA4 defined, among others, the concept of a Media Service Enabler (MSE) in *3GPP TR 26.857* [1] as shown in Figure 1. MSEs address the support of third-party applications to make use of advanced functionalities provided by the 5G System, combined with additional well-defined client and network functionalities for media services. MSEs mimic what is referred to in implementations and deployments as Software Development Kit (SDK), i.e. a set of packaged functions that are usable by applications through well-defined APIs. Properties assigned to MSEs is the abstraction of complex 5G System and media functions into developer-friendly APIs, reducing the complexity of developing applications. It is also expected that MSE create a set of testable functions, whereby testing and conformance may be addressed outside 3GPP, for example by a Market Representation Partner (MRP) such as 5G-MAG or by an industry forum. MSEs are supported by developer-friendly guidelines and examples to make use of the set of functionalities provided by an MSE.

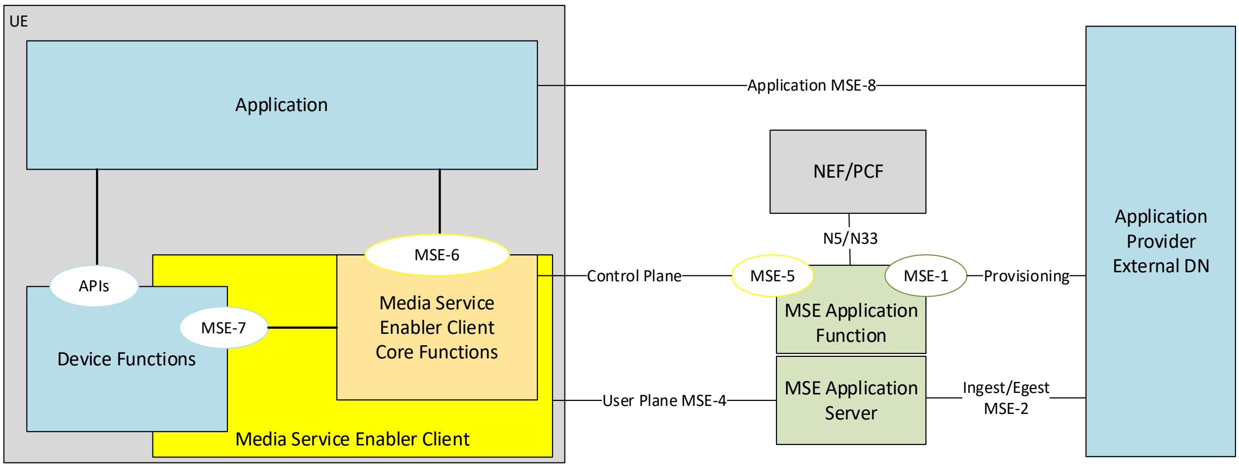

Figure 1 Media Service Enabler – Basic Architecture

Key Work topics

In the following, a few key work topics are addressed in 3GPP SA4 are introduced.

User Services and user service components: 3GPP SA4 addresses the establishment of user services such as telephony, multi-party conferencing, streaming, multicast and broadcast distributions, etc. Service layers are developed to establish connectivity, access to relevant system resources and establishment of end-to-end services. In certain cases, only components are provided that may be integrated into end-to-end third-party services. Examples for such user services are 5G Media Streaming (see TS 26.501 [2]), MBS User Services (see TS 26.502 [3]) and Real-time Communication Services (see TS 26.506 [4]). User Services are supported by user-plane content delivery protocols such as DASH/HLS/CMAF (see TS 26.511 [5], TS 26.512 [6] and TS 26.247 [7]) and RTP (see TS 26.113 [8], TS 26.114 [9] and TS 26.522 [10]) with their adaptation to 3GPP service needs. Also recently, Messaging Services based on renewed efforts on 5G Messaging and interoperability needs across proprietary systems have been upgraded to support latest media technologies in TS 26.143.

Integration with 5G System and QoS: 5G provides elaborate Quality-of-Service (QoS) functionalities, for example the ability to assign certain bitrate, delay and reliability provisioning to ensure sufficiently high quality for the service. The mapping of modern media services, including 3GPP defined services as well as third-party services, to these QoS system is addressed in core specifications developed by 3GPP SA4. Examples include mapping to SDP parameters to QoS service parameters (see TS 26.114 [9]) and dynamic policies as defined in TS 26.501 [2] and TS 26.506 [4].

Codecs and Formats: 3GPP provides requirements and guidelines for media formats and codecs in order to provide interoperable media services. For speech and audio, 3GPP SA4 has a long history to develop codecs, but for video, 3GPP primarily relies on technologies developed externally. 3GPP SA4 provides profiles for codecs and formats that match service requirements, including codecs for telephony services (see TS 26.117 [11]), for TV Services (see TS 26.116 [12]), for streaming services (see TS 26.511 [5] and TS26.117 [11]), for VR services (see TS 26.118 [13]) and for AR/XR (see TS 26.119 [14]).

Metrics and QoE: 3GPP SA4 is also responsible to support the definition and monitoring of Quality-of-Experience (QoE). For this purpose, metrics collection is addressed in 3GPP SA4 specifications such that the information may be used for QoE evaluation in 3GPP service operation or provided to external providers. Frameworks for metrics collection are defined in TS 26.531 [16] and TS 26.532 [17] and instantiated in different services such as 5G Media Streaming in TS 26.501. Metrics are for example defined for streaming services in TS 26.247 or for VR services in TS 26.118.

Distributed compute for media: With the advance of cloud and edge computing, user services more and more rely on distributed processing, partially done on the UE and partially done in the network. Examples include the user of Media Resource Functions (MRF) for multi-party conferencing, edge and cloud processing for 5G Media Streaming as defined in TS 26.501, split rendering as evaluated in TR 26.928 [18] for XR, tethering architectures including AR Glasses, phones, edge and cloud as evaluated in TR 26.806 [19] or split inference for distributed AI processing as currently evaluated in TR 26.927 [21]. Distributed compute requires understanding of the media formats to be exchanged, but also typically needs specific network support in terms of QoS.

In the following, specific selected activities completed in Rel-18 and in progress in Rel-19 are introduced in more details.

EXAMPLE 1 - USER SERVICES: EXTENSIONS TO 5G MEDIA STREAMING

In 3GPP Release 16, the core principles of the 5G System had been extended for media services by defining the 5G Media Streaming (5GMS) architecture. 5GMS is built on the idea of enabling third-party media distribution beyond the MNO (Mobile Network Operator) and the 5G network acting not only as a bit pipe but to provide technical and commercial opportunities for collaboration. Challenges for pure over-the-top distribution include: (i) Quality of Experience issues (rebuffering and stall events) that are associated with the operator in the end user’s mind, (ii) obscuring of traffic by end-to-end encryption using HTTPS, (iii) unidentified content eating into users’ data caps, and (iv) the increasing demand for higher quality media, new formats, as well as new immersive and interactive experiences.

The 5G Media Streaming architecture is documented in 3GPP TS 26.501 [2] and shown in a generalized way in Figure 2. The framework is aligned with modern over-the-top media distribution practices. The specifications support MNOs and third-party media services to easily access 5G System and 5G Media Streaming features. Specifications 3GPP TS 26.510 [22], TS 26.511 and TS 26.512 an instantiation of 5G Media Streaming for one or a small subset of recommended technologies including codecs, formats, protocols, APIs and other functionalities.

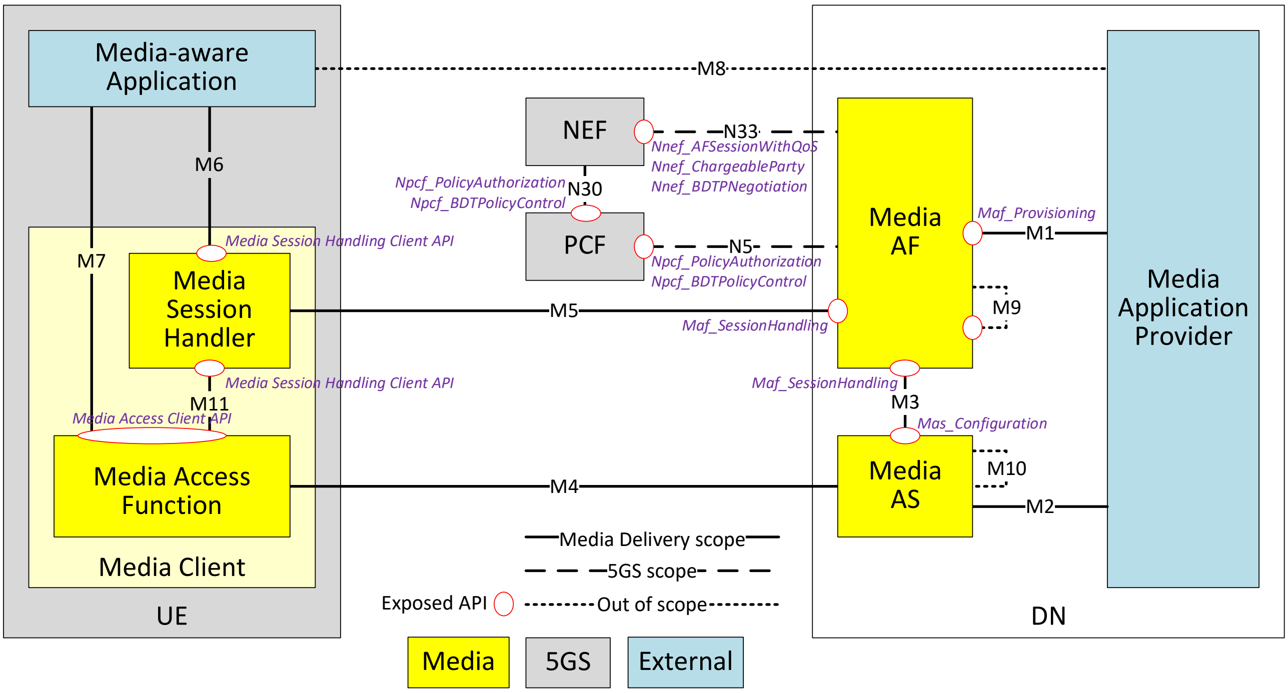

Figure 2 3GPP Media Delivery Architecture used and instantiated by 5G Media Streaming

In the course of extensive studies documented in TR26.804 [23], a set of stage-2 and stage-3 extensions to 5G Media Streaming had been completed in Rel-18 to update TS 26.501 and TS 26.512. TS 26.501 defines the 5GMS architecture, call flows, and procedures. TS 26.512 defines the 5G Media Streaming protocols. Furthermore, when producing to TS 26.506 (5G Real-time Media Communication Architecture), many commonalities between streaming and real-time delivery were identified. Many of the features developed primarily for "5G Media Streaming"-based Session Handling are applicable to different service scenarios. Hence, a harmonized Media Session Handling taking into account stage-2 from 5G Media Streaming (TS 26.501) and Real-time Delivery (TS 26.506) was developed. Based on the updates in TS 26.501, the new specification TS 26.506 and the result of an extensive study in Rel-17, with outcome documented in TR 26.804, the Rel-18 work item addresses in total 12 topics that are added to 3GPP specifications, namely

1) Consistent support for uplink streaming – this functionality is needed for media production services, but also for contributing user generated media to operator and third-party services. Different collaboration scenarios between the MNO and the service provider are addressed in an updated specification – bringing uplink streaming on par with downlink streaming.
2) Support for end-to-end low latency live streaming and combining this with dynamic policies is of crucial support to meet the TV industry needs in order to bring 5G on par with other distribution means. Live TV services are characterized by at least the following aspects: (1) scalability (in terms of concurrent users), (2) consistent quality, (3) high bandwidth requirements, (4) target latency constraints, and advanced TV Experiences.
3) Support for 5GMS over MBS and 5GMS hybrid services
4) Support for multiple media service entry points: Based on development in CTA WAVE in context of DASH/HLS interop specification in CTA-5005, it was identified that the same media service may be available in multiple formats, and 5GMS media is expected to offer the same service with different entry points in order for the device to select the service based on its capabilities, for example using a DASH player, or and HLS player, or using a progressive download system

5) Extensions to 5GMS protocols to support traffic identification.
6) Addition of necessary parameter extensions to the M1, M5, and M6 reference points to provide access to Background data transfer (BDT). BDT enables mobile network operators to incentivize the offloading of traffic into off-peak hours with the aim of reducing network congestion during peak hours. Applications may register their interest in receiving content using BDT.
7) Specification of the usage of Oauth 2.0 (according to the SA3 guidelines) for 5GMS protocols
8) Specifications for the 3GPP Service Handler and URL including the necessary functions on UE and device to support automatic launch of 5G System services in the context of 5G Media Streaming. It was identified that it is essential to provide a URL form for media services in order to automatically launch client and network functions without disrupting existing app workflows.
9) Specification of a RESTful API at reference point M3, for the configuration of 5GMS AS instances by 5GMS AF. M3 was undefined until now, but based on deployment experience the need for more details at this reference point was identified.
10) Specification of data types for data reporting of ANBR-based Network Assistance invocations and for exposure of events relating to invocation of AF-based and ANBR-based Network Assistance.
11) Enhancements based on feedback from 5G-MAG Reference Tools developments. The support of 5G-MAG Reference Tools created essential inputs to the spec development and continuous exchange will result in additional updates.

For Rel-19, an Advanced Media Delivery study is initiated in 3GPP with broad support of different industry players. While the Rel-18 work was primarily focussing on session handling features, it was recognized that media streaming delivery protocols are continuously enhanced to address deployment challenges. The primary focus of Rel-19 study and work is the delivery of segmented media objects in the media plane, i.e. at reference points M2, M4 and M7 of the Media Delivery architecture. The feasibility study also addresses topics related to MBS/MBMS, which initially may be considered orthogonal to user plane aspects of segmented media delivery. However, generally MBS/MBMS and unicast user plane issues are preferably handled jointly because MBS/MBMS is considered a “transparent” transport pipe. Many of the functionalities on the M2/M4/M7 media plane are expected to be available as well for MBS/MBMS-delivered media. Among others, the following aspects are under study:

1) Common Client Metadata: CTA WAVE has developed CTA-5004: Web Application Video Ecosystem Common-Media-Client-Data (CMCD) [24]. It is worthwhile to study the benefits of integrating commonly supported metrics and client data reporting in 5GMS workflows.
2) Common Server-and Network-Assisted Streaming: MPEG-DASH supports Server and Network Assisted DASH (SAND) [25]. Certain profiles of SAND were adopted in TS 26.247, but the industry has generalized the concepts in SAND in efforts such as Content Steering (see ETSI TS 103 998 [26]), Web Application Video Ecosystem (WAVE) specification for Common Media Server Data (CMSD) [27], or Addressable Resource Index (ARI) Tracks in MPEG [28]. The study and integration of these technologies into the Media Delivery System and MBS/MBMS workflows is of significant interest, in particular also in combination with existing QoS mechanisms.
3) Multi-CDN and Multi-Access Media Delivery: Content distributors often use multiple Content Delivery Networks (CDNs) or multiple access networks to distribute their

content to end-users. As an example, they may upload a copy of their catalogue to each CDN, or more commonly have all CDNs pull the content from a common origin.

4) Modem Usage Optimized Media Streaming: Enhancements to Background Data Transfer to support preload as well as functionality of what is defined in W3C Managed Media Source Extension to minimize active network connections are relevant topics to study with the aim of limiting battery consumption in the UE resulting from media delivery.
5) DRM and Conditional Access: DRM and Conditional Access are commonly used by third-party streaming services. However, in case streaming is done through MBS or MBMS, a more careful management of the keys needs to be checked. Scalability of key delivery is an issue.
6) In-session Unicast Repair for MBS Object Distribution: For live and low-latency live services using the Object Distribution Method in MBS, in certain cases the transmission of an object is not successful. In this case, unicast repair for individual MBS Clients can improve the service quality.
7) MBS User Service and Delivery Protocols for eMBMS: In order for MBMS and LTE-based 5G broadcast to leverage MBS User Service technologies, a study is warranted to identify the gaps to fully support this functionality.
8) Selected MBMS Functionalities not supported in MBS: In completing TS 26.502 and TS 26.517 [29], it is obvious that only a subset of the MBMS functionalities is supported in Rel-17. While many MBMS functionalities are likely not important to be supported for MBS, a systematic analysis of MBMS User Services features and their potential relevance for MBS should be completed and recommendations made on which ones to migrate to MBS User Services specifications and how best to achieve this.
9) DASH/HLS Interoperability: DASH/HLS interoperability is a key issue to support highly scalable distribution systems for CDN-based and MBS/MBMS distribution. Offering common CMAF segments that can be consumed by both DASH and HLS media players promises to address these issues.
10) Further harmonization of RTC and Streaming for Advanced Media Delivery: With the creation of TS 26.510 in Rel-18, Media delivery across the 5G Media Streaming (5GMS) System and the Real-Time media Communication (RTC) System was harmonized. Study of further harmonization is encouraged to fully implement common Media Delivery functions.
11) Improved QoS support: In Rel-18, SA2 has defined a number of new features in the 5G System, especially in the PCF, from which media delivery may benefit. Examples documented in TS 23.501 [30] include the use of Explicit Congestion Notification (ECN) to support Low Latency, Low Loss, Scalable Throughput (L4S) services (clause 5.37.3), PDU Set handling (clause 5.37.5) and QoS Monitoring (clause 5.45), and there are likely others. The impact and usefulness of selected features is preferably studied.
12) Impacts and opportunities of QUIC for segmented content delivery: Since the finalisation of the QUIC protocol by the IETF in May 2021, there has been significant deployments of QUIC driven by the usage of HTTP/3 [31] for streaming services. In the IETF, the working group on Media Over QUIC (MOQ) is working towards an extensible protocol for publishing media for ingest and distribution.

EXAMPLE 2 - CODECS AND MEDIA FORMATS FOR XR

Often seen as the key component of the soon to arrive Metaverse, Extended reality (XR) refers to all real-and-virtual combined environments and associated human-machine interactions generated by computer technology and wearables. It includes representative forms such as augmented reality (AR), mixed reality (MR), and virtual reality (VR) and the areas interpolated among them. There has been the emergence of a wide range of AR capable devices beyond smartphones such as AR Glasses and Mixed Reality Head-Mounted Displays (HMD). For service providers and application developers, addressing this wide spectrum of devices and their capabilities is a challenge and even prevents the wider deployment of AR services. In Rel-18, the 3GPP TS 26.119 specification was developed with the aim at creating a higher degree of interoperability between application providers, content creator and manufacturers by consolidating the media capabilities of AR-capable devices in handful of well-defined device categories.

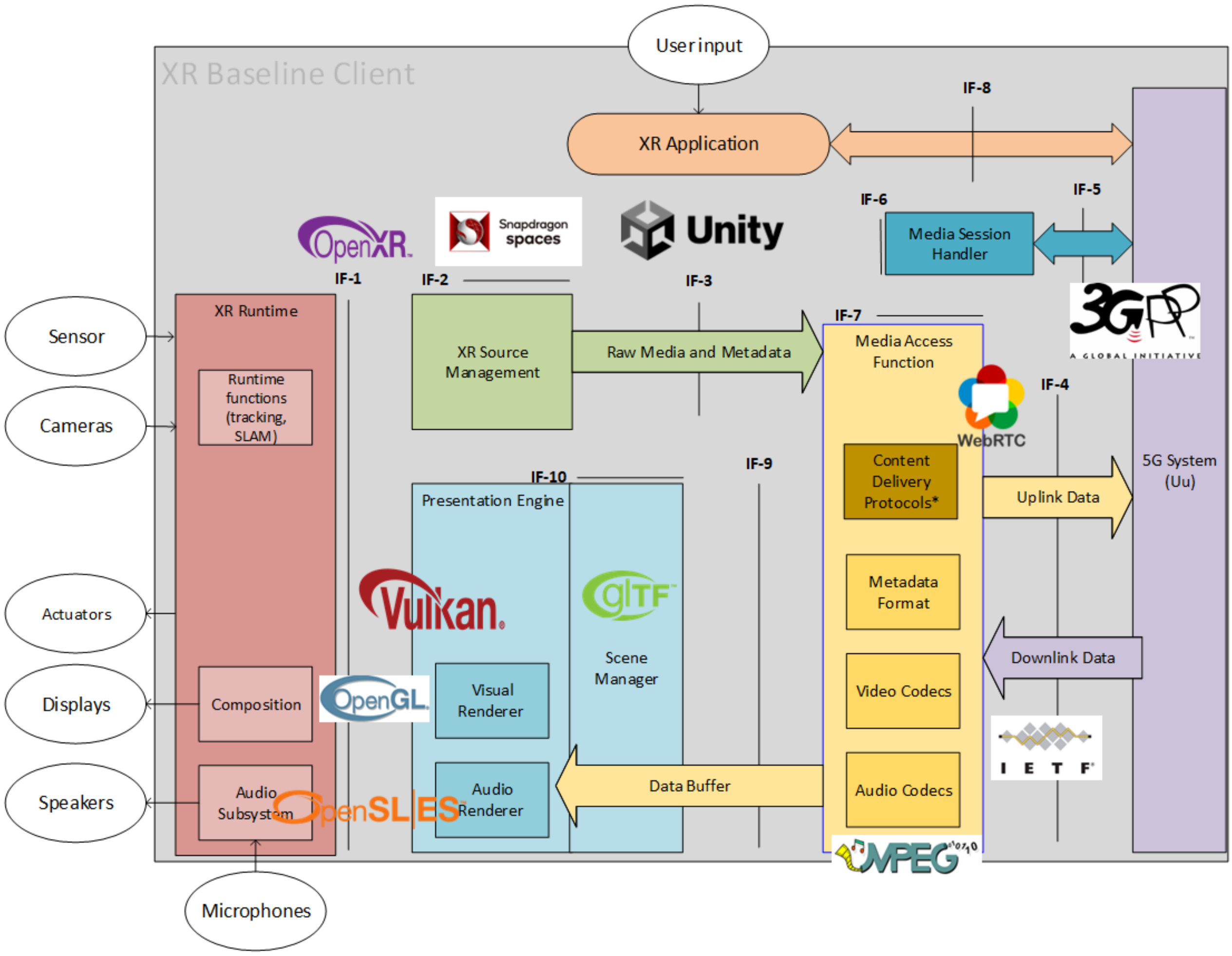

Figure 3 XR Baseline Client with example technologies used for different building blocks

This XR client in UE terminal architecture as shown in Figure 3 is in line with industry standards such as the OpenXR specification defined by the Khronos Group which addresses jointly AR, MR (Mixed Reality) and VR (Virtual Reality) experiences. At the top of the architecture seats the Application which is responsible for orchestrating the various components to provide the AR experience. The specification involves a set of profiles inherited from Khronos, MPEG, IETF and other 3GPP specs, combining the functions to access, decode and render immersive media experiences. The XR Runtime is responsible for the final audio and visual rendering based on the views supplied by the Presentation Engine. The XR Runtime is typically defined by XR industry framework such as OpenXR. The XR Runtime also exposes sensor data such as the pose prediction which can be retrieved and packaged before transport by the XR Source Management. Lastly, the Media Access Function is the media powerhouse to process the delivered scene description and to decompress the audio and video data received via the 5G System which composes the immersive experience.

Based on the XR client in UE terminal architecture, TS 26.119 defines four categories of devices targeting various types of form factors. (i) Thin AR glasses, (i) AR glasses, (iii) XR phone and (iv) XR Head Mounted Display (HMD), addressing different power and complexity constraints. For each device type, TS 26.119 specifies the mandatory and optional media capabilities to be supported by the UE. Those media capabilities pertain to audio, video, scene processing and XR systems capabilities and are defined in terms of supports of audio codecs (EVS, IVAS and AAC-ELDv2), video codecs (AVC and HEVC), scene description formats (glTF 2.0 and its extension in MPEG-I Scene Description). In addition, common metadata are defined in TS 26.119 that are predicted pose information, the action object representing actions performed by a user of an AR application and the available visualization space object representing a 3D space within the user's real-word space that is suitable for rendering virtual objects.

In Release-19, normative work and studies are ongoing to extend the video capabilities including Beyond 2D formats. In particular, it was identified to have a well-defined capability for stereo video using MV-HEVC based encoding to provide interoperability between the iOS and Apple Vision Pro ecosystems with the Android and Meta devices. 3GPP SA4, together with MPEG, is currently developing a streaming and messaging format to address this need. Beyond this normative work, a study is initiated to evaluate Beyond 2D video formats for the exchange within 3GPP networks. Different scenarios are developed addressing the exchange of Beyond 2D formats in messaging services, streaming beyond 2D formats originating from professional produced live shows, user-generated content as well computer-generated or AI enhanced content. Multiview-formats including depth, and other new representation formats are expected to be investigated for their suitablity to be used within 5G or 6G services.

EXAMPLE 3 - DISTRIBUTE COMPUTE - SPLIT RENDERING MSE

Advances in computer graphics and machine learning have enabled a wide range of new experiences and applications for users. From remote gaming to autonomous driving, complex and power-hungry processing is required to achieve the desired user experience. As an example, recent games are reverting to ray tracing and global illumination to offer a physically based rendering solution that will mimic the physical world accurately and immerse the user in the game scene.

The 5G system offers several capabilities that would pave the way for making these demanding media services available to all users, independently of their end device capabilities. For example, access to edge computing allows the rendering of complex 3D scenes in powerful edge servers and then displaying them on the user's end devices. This functionality is denoted as Split Rendering. It is also supported by QoS allocation to ensure that the operation takes place smoothly. The applications of split rendering may vary widely and its usage for AR has been proven required for lightweight devices. Therefore, a split rendering media service enabler to support Edge-assisted XR devices was defined as part of the 3GPP release 18. Other XR applications, such as immersive 6DoF (6 degrees of freedom) streaming and online gaming are also set to benefit from a split rendering enabler to deliver highly immersive experiences to their users.

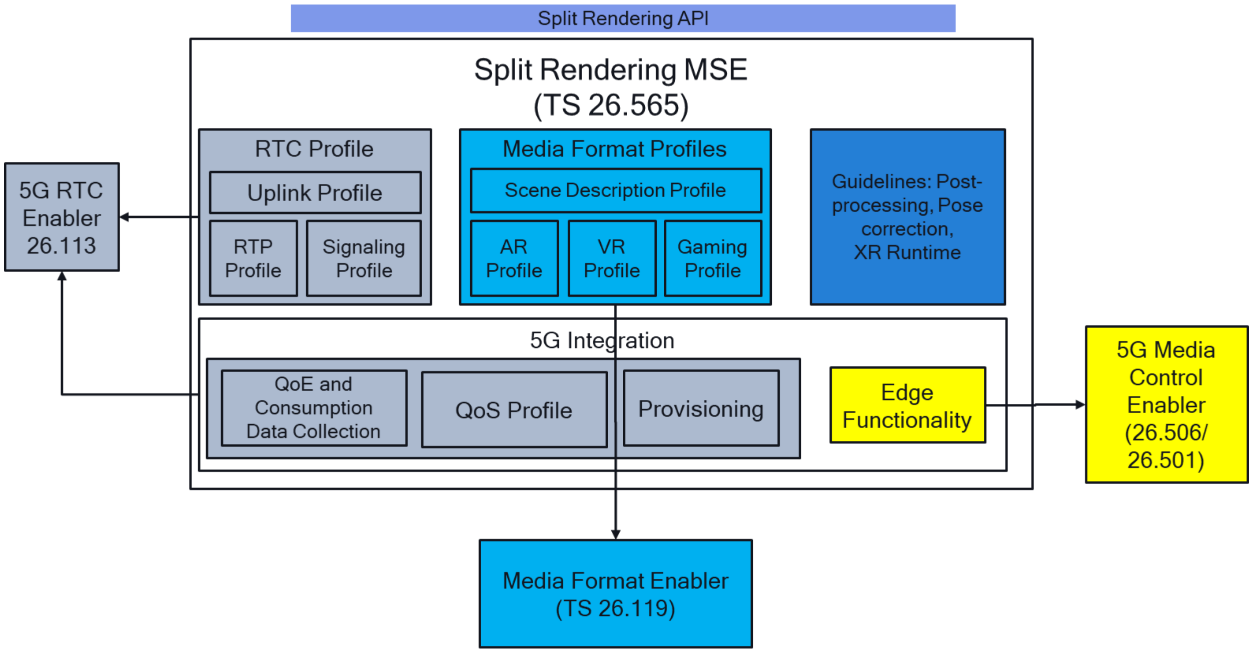

Figure 4 Split Rendering Media Service Enabler Protocol Stack

TS 26.565 [32] defines a Media Service Enabler for split rendering. The split rendering MSE allows developers to easily create applications that are able to automatically delegate rendering to the edge. It leverages several enablers to simplify the deployment of this enabler by abstracting details of edge discovery, format negotiation, connection establishment, QoS policy allocation, etc. The specification also defines the media and metadata formats for split rendering. It provides guidelines to the split rendering server on how to encode the rendered media. In its current version, TS26.565 defines several split rendering profiles that support different types of devices and different split strategies. Figure 4 provides the Split Rendering Media Service Enabler protocol stack that makes use of Media Delivery functions, formats defined in TS 26.119, protocols defined in TS 26.113 as well as functionalities to establish split rendering sessions.

MORE HIGHLIGHTS IN SA4 REL-18 AND REL-19

A few more highlights completed in Rel-18 and ongoing in Rel-19 are provided as follows:

- New 3GPP codec for Immersive Voice and Audio Services (IVAS) – Rel-18. The overall objective of this work item is to develop a single general-purpose audio codec for immersive 4G and 5G services and applications including the XR use cases.
- Terminal Audio quality performance and Test methods for Immersive Audio Services (ATIAS) – Rel-18: The overall objective of this work item is to develop a set of test specifications for objective characterization of terminals for 3GPP immersive services.
- Immersive Audio for Split Rendering Scenarios (ISAR) – Rel-18: The overall objective of this work item is to develop solutions for immersive binaural audio on head-tracked devices that are compatible with the envisaged split architectures.
- 5G-Advanced media PROfiles for MessagIng Services (PROMISE) – Rel-18: The purpose of this Work Item is to specify SMS/MMS/RCS/Messaging formats and codecs that GSMA and other organizations or application vendors can then reference and/or profile to improve messaging service quality and interoperability. It includes messages with features such as advanced audio, HEIF-based images, and 3D scenes and assets.

- Feasibility Study on Artificial Intelligence (AI) and Machine Learning (ML) for Media (FS_AI4Media) – Rel-19: The objectives of the study item are primarily to identify relevant interoperability requirements and implementation constraints of AI/ML in 5G media services.
- Feasibility Study on Avatars for Real-Time Communication (FS_AVATAR) – Rel-19: Based on relevant TR 22.856 use cases, the study item aims, among other objectives, to collect and document Avatar animation and representation approaches, document the requirements for an interoperable base Avatar format.
- Immersive Real-time Communication for WebRTC (iRTCW) – Rel-18: Define a non-vertical/modularized protocol stack, functional components, I/Os formats, for iRTC clients to support WebRTC-based real-time transport of media over 5G; Identify the required architecture for radio access network QoS realization over 5G systems. Identify the minimum information / elements in the C/U-Plane signal to establish media sessions with appropriate QoS for WebRTC-based applications. Document informative examples of iRTC operations to assist implementers.
- IMS-based AR Conversational Services (IBACS) – Rel-18: The objective of this work item is to create a new specification for IMS-based AR conversational services. The features for RTP-based real-time communication, which can be used by IMS and non-IMS (AR) conversational services, will be specified in another new specification (as part of the 5G_RTP work). The relevant features and functional components specified for MTSI in TS 26.114 will be referenced and/or enhanced, if required.
- 5G Real-time Transport Protocols (5G_RTP) – Rel-18: The objective of this work item is to specify functionalities of RTP to improve support for traditional and immersive real-time services and enablers. To develop a commercially relevant set of functionalities that only include technologies that are either commercially relevant or deployed or demonstrate clear performance or relevant functionality that justifies introducing additional implementation or interoperability complexity. The work is extended in a phase 2 in Rel-19.

REFERENCE IMPLEMENTATIONS SUPPORTING VALIDATION AND TESTING

The 5G-MAG Reference Tools (see developer.5g-mag.com) are making available reference implementations with a focus on the specifications led by 3GPP SA4. The set of reference tools and projects include, among others:

- 5G Media Streaming server- and client-side Android-based components. Currently implemented features include APIs and procedures for QoE metrics and consumption reporting, content hosting, and network assistance.
- Delivery Protocols for eMBMS and end-to-end chain including server side (except BM-SC) and Android-based client components.
- Initial implementation of MBS support in the 5G Core.
- XR content compliant with scene description formats (glTF 2.0 and MPEG-I Scene Description extensions) and reference Unity player.

CONCLUSIONS

This paper presents an overview of the latest developments on the methodologies and work topics in 3GPP with respect to media. Focus is on advanced use cases that address immersive as well as AI-based media, but also addresses continuous improvements of operator and third-party services. With 6G in the horizon, it is expected that the ever-attractive media services will play a central role in the definition of new capabilities, including the ability of functional improvements (APIs, developer-friendly, cost-efficient, sustainable) as well as new use cases including new media addressing XR and the Metaverse.

ACKNOWLEDGEMENTS

The authors would like to thank all the members involved in the 3GPP SA4 standardization activity. We would also like to thank the International Broadcasting Convention for permission to publish this paper.

REFERENCES

1. 3GPP TR 26.857 – 5G Media Service Enablers
2. 3GPP TS 26.501 – 5G Media Streaming (5GMS); General description and architecture
3. 3GPP TS 26.502 – 5G multicast-broadcast services; User service architecture
4. 3GPP TS 26.506 – 5G Real-time Media Communication Architecture (Stage 2)
5. 3GPP TS 26.511 – 5G Media Streaming (5GMS); Profiles, codecs and formats
6. 3GPP TS 26.512 – 5G Media Streaming (5GMS); Protocols
7. 3GPP TS 26.247 – Transparent end-to-end Packet-switched Streaming Service (PSS); Progressive Download and Dynamic Adaptive Streaming over HTTP (3GP-DASH)
8. 3GPP TS 26.113 – Real-Time Media Communication; Protocols and APIs
9. 3GPP TS 26.114 – IP Multimedia Subsystem (IMS); Multimedia telephony; Media handling and interaction
10. 3GPP TS 26.522 – 5G Real-time Media Transport Protocol Configurations
11. 3GPP TS 26.117 – 5G Media Streaming (5GMS); Speech and audio profiles
12. 3GPP TS 26.116 – Television (TV) over 3GPP services; Video profiles
13. 3GPP TS 26.118 – Virtual Reality (VR) profiles for streaming applications
14. 3GPP TS 26.119 – Media Capabilities for Augmented Reality
15. 3GPP TS 26.265 – Media Delivery: Video Capabilities and Operating Points
16. 3GPP TS 26.531 – Data Collection and Reporting; General Description and Architecture
17. 3GPP TS 26.532 – Data Collection and Reporting; Protocols and Formats
18. 3GPP TR 26.928 – Extended Reality (XR) in 5G
19. 3GPP TR 26.806 – Study on Tethering AR Glasses – Architectures, QoS and Media Aspects
20. 3GPP TS 26.143 – Messaging Media profiles
21. 3GPP TR 26.927 – Study on Artificial Intelligence and Machine learning in 5G media services
22. 3GPP TS 26.510 – Media delivery; interactions and APIs for provisioning and media session handling
23. 3GPP TR 26.804 – Study on 5G media streaming extensions
24. CTA-5004 – Web Application Video Ecosystem - Common Media Client Data
25. ISO/IEC 23009-5 – Information technology — Dynamic adaptive streaming over HTTP (DASH)Part 5: Server and network assisted DASH (SAND)

26.ETSI TS 103 998 – DASH-IF: Content Steering for DASH

27.CTA-5006 – Web Application Video Ecosystem - Common Media Server Data

28.ISO/IEC 23009-1 – Information technology — Dynamic adaptive streaming over HTTP (DASH) Part 1: Media presentation description and segment formats

29.3GPP TS 26.517 – 5G Multicast-Broadcast User Services; Protocols and Formats

30.3GPP TS 23.501 – System architecture for the 5G System (5GS)

31.IETF RFC 9114 – HTTP/3

32.3GPP TS 26.565 – Split Rendering Media Service Enabler