GTI
XR Network Technology White Paper

GTI XR Network Technology White Paper

Version:	**v_1.0.0**
Deliverable Type	**□ Procedural Document** **☑ Working Document**
Confidential Level	**□ Open to GTI Operator Members** **☑ Open to GTI Partners** **□ Open to Public**
Program	5G Technology and Product
Working Group	
Project	Technology Evolution
Task	5G-A Technology
Source members	China Mobile, HUAWEI, ZTE, Ericsson, Qualcomm, MTK
Support members	Comba Network, Sercomm
Editor	Lei Cao(CMCC), Luting Kong(CMCC), Chenguang Jin(CMCC), Hongbiao Zhang(CMCC), Tianming Jiang(CMCC), Xin Li (CMCC), Jing Gao(CMCC), Bo Wang(CMCC), Lu Li(CMCC), Wei Wen(HUAWEI), Pin Jiang(HUAWEI), Yi Qin(HUAWEI), Xintai Wang(ZTE), Ting Li(ZTE), Meng Lai(Ericsson), Liyu Liu(Ericsson), Bo Chen (Qualcomm), Zhichao Zhou(Qualcomm), Mingyuan Zheng(MTK), Peng Yang(MTK)
Last Edit Date	2023-06-22
Approval Date	2023-06-25

Document History

Date	Meeting #	Version #	Revision Contents
2023.6.22	GTI#37	V1.0	Initial Version

Table of Contents

Introduction

Extended reality (XR) is currently a hot topic around the world. In 2023, China rolled out a plan for the overall layout plan for the country's digital development, which first mentioned "immersive service experience". Extended reality will play an important role in the construction of digital society.

5G is the key infrastructure for the digital transformation of the economy and society. This white paper aims to explore the key capabilities of 5G networks, devices and services to satisfy the large bandwidth and low latency requirements simultaneously for XR to achieve a new immersive service experience. In this white paper, the first chapter makes an overall introduction to XR technology and industrial development. The second chapter analyzes the requirements of XR service and the challenges brought by XR service to 5G network. The third chapter expounds the key technical solutions for XR including network awareness service, service adaptation network, Connectivity Improvement , rendering capability collaboration, and user experience evaluation. Finally, the application scenario and business model of XR is discussed.

1. Overview

Extended Reality (XR) is the umbrella term for all immersive technologies including VR, AR, and MR.

Virtual Reality (VR) refers to bringing the user's perception into a totally virtual world by wearing devices. Users could experience the most real feelings in the virtual reality world. The authenticity of virtual world is difficult to distinguish from the real world. VR realizes the immersive interaction between people and the virtual world.

Augmented Reality (AR) refers to that the virtual information generated by computers is effectively applied in the real world. The interaction between people and the real world is more direct and natural with AR. With the continuous development of AR technology and the continuous evolution of 5G mobile communication network, AR is expected to truly eliminate the sense of boundary between virtual and reality.

Mixed Reality (MR) refers to the combination of VR and AR technology. MR Has both VR and AR functions and advantages. The real world, virtual world and digital information could be combined to achieve the real world and virtual world integration interaction.

1.1. The development history and trend of XR

XR headset devices already had some mature applications in military and other fields before 2012, but limited by the cost and technical maturity of XR headset devices, most of the early commercial attempts of XR devices have failed. From 2012 to 2014, the XR industry entered a period of development. Google released Google Glass, marking the advent of the concept of consumer AR devices. Facebook successfully acquired Oculus, a VR head display manufacturer, and launched Oculus rift VR headset devices. From 2014 to 2017, the XR industry entered a another development boom. Hardware companies such as Microsoft, HTC and Samsung launched consumer products one after another. After the official release of the AR game Pok Pokémon Go, it quickly became popular all over the world. However, the technology of XR headset device is still immature. For example, the screen clarity is low, the user experience is not good enough and the production cost of application is high which leads to the low purchase intention of users. From 2017 to 2019, the overall development of the XR consumer market slowed down. From 2019 to now, With the continuous iteration of XR headset technology and the continuous precipitation of applications,, VR has begun to exert its strength on the content side of games and social contact. Qualcomm launched the first XR chipset platform XR2 which also supports 5G, industry benchmark products such as Oculus Quest 2 and Pico Neo3 came out one after another, and the VR industry began to recover. In 2021, VR shipments exceeded 10 million for the first time,. In June 2023, Apple released its first MR device Apple Vision Pro. Apple Vision Pro introduces a new interaction system, which gets rid of external controllers and interacts with eyes, gestures and voice. It is expected to explode the industry again and lead new changes in the XR field.

In future, with the continuous development of communication technology, AI computing, edge cloud and other infrastructure construction, XR games, entertainment, social, education and other content will continue to enrich. XR, as the early entrance of the meta-universe, will bring revolutionary changes and innovative experience to people's lifestyle and working mode.

1.2. Observation of XR Industry

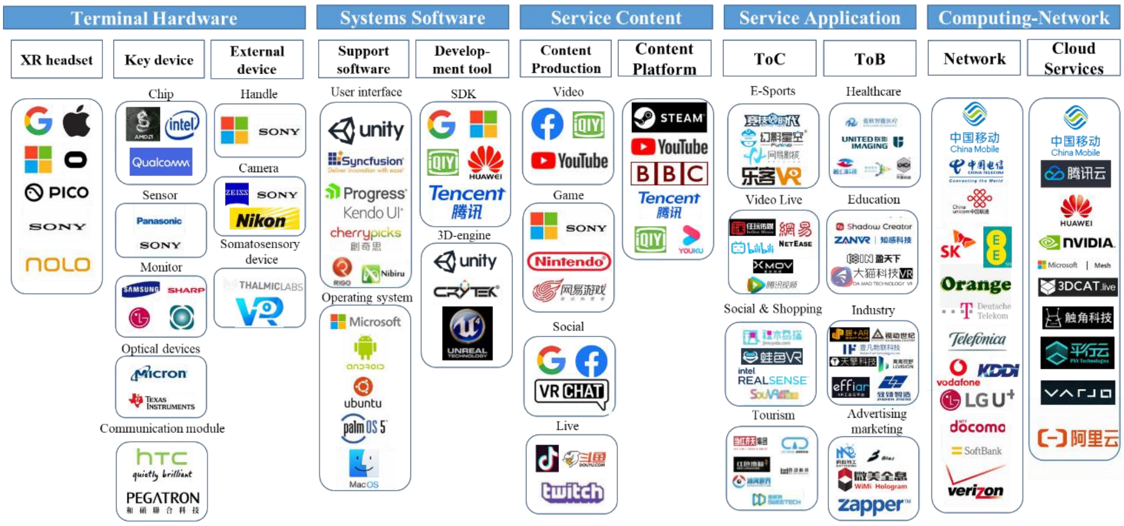

FIG.1 XR industry chain scan

XR industry chain involves applications, system software, network services, chipsets, devices and other aspects. In the industry chain, there are still pain points and difficulties that restrict the high-quality and large-scale development of XR, which need to be overcome jointly by the industry.

- Content and application needs to be enriched. VR applications focuses on video, games, and social scenes, while content in other areas is less involved. Due to the high production cost, long production cycle and the small size of the XR market, large factories are still in a wait-and-see state, resulting in a low proportion of VR content. For example, on the global mainstream game platform Steam, VR content accounts for less than 5%; AR content ecology has not yet been fully built, and most of the current AR applications are transplanted to mobile 2D applications, which is difficult to stimulate consumers' desire to buy.
- Device needs to be lightweight. Limited by the optical scheme and battery, the current XR devices is still relatively heavy, affecting the user experience. The Fresnel lens optical solution used in VR/MR Headsets has a long imaging distance and a thick optical component, Pancake solutions have been gradually adopted in the new product to help reduce the imaging distance. The mainstream headset battery life is about 2 to 3 hours due to the high power consumption of the screen, high-performance chipsets. Some new products remove the battery from the headset by means of using external battery to reduce the weight of the

headset. In addition, some newly released headsets with 5G capabilities are developed based on 5G modules with low integration. It is necessary to promote the launch of highly integrated products based on 5G chipsets.

- Network capability needs to be further improved. XR brings new requirements for large bandwidth, low latency, network capacity and reliability. It is necessary to research the connection enhancement technology. Details will be discussed in the following chapters.

2. The Challenges brought by XR to 5G Networks

2.1. XR service requirements

XR services are diversified, and are divided into four categories: high-fidelity strong interaction, high-fidelity weak interaction, low-fidelity strong interaction, and low-fidelity weak interaction according to different requirements on network bandwidth and delay.

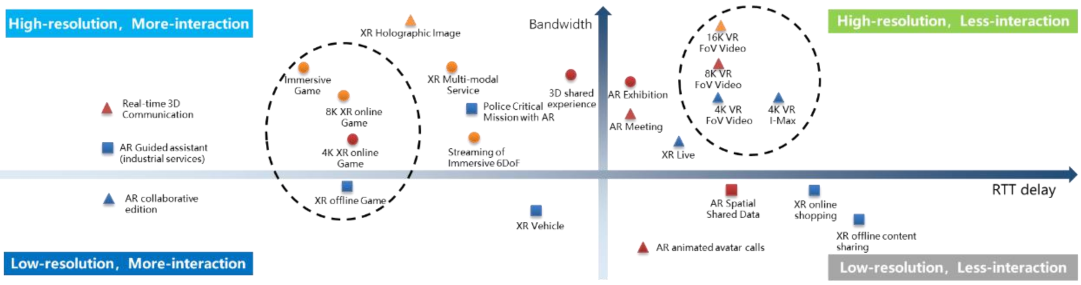

FIG.2 Classification of XR services based on latency and bandwidth requirements

- High-fidelity strong interaction services which demand for large bandwidth and low latency simultaneously. Typical services include 4K XR online games, 8K immersive XR online games, multi-sensory virtual collaboration/social/games. For example, 4K interactive games require an average speed of around 50Mbps and an end-to-end network latency of <20 ms [1].
- High-fidelity weak interaction services which demand for large bandwidth,,Typical services include VR FOV video, VR giant screen theater, AR movie viewing, etc. For example, the bandwidth requirements of 8K FOV video is more than 100Mbps [1]
- Low-fidelity strong interaction services demand for low latency. Typical services include vehicle XR services and AR multi-party collaboration. For example, AR multi-party collaboration requires an end-to-end network delay <15 ms [1].
- Low-fidelity weak interaction services which the requiremtns for bandwidth and latency are lower than the above three types of services. Typical services include AR online shopping.

2.2. Key Network Metrics for XR services

With the evolution of XR from "virtual-real connection" to "virtual-real symbionism", data flow expands from audio-visual data to audio-visual-tactile data, image resolution evolves from 4K to 8K, user experience develops from semi-immersive to fully immersive, application scenarios develop from indoor to outdoor, and the number of users increases with the popularization of XR application, cellular networks will face challenges of data rate, latency, capacity and mobility.

2.2.1. Latency

Take VR cloud game as an example, the MTP latency could be divided into seven parts: action capture delay → network transmission delay (uplink) → computing and rendering delay at cloud side → coding delay → network transmission delay (downlink) → decoding delay → screen rendering display delay.

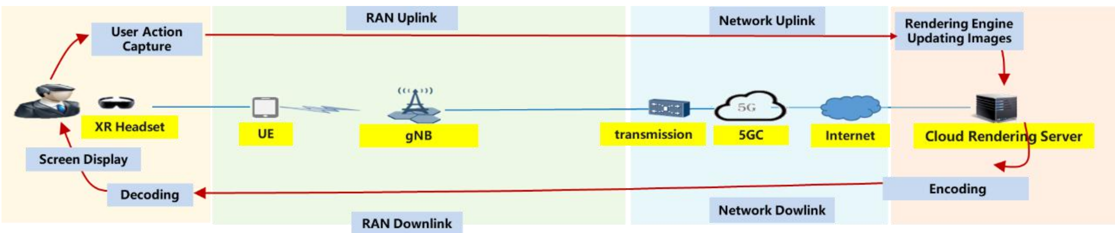

FIG.3 End-to-end delay decomposition diagram

Theoretically, in order to ensure a good XR user experience and no symptoms such as dizziness, the Motion to Photons (MTP) is required to be about 20ms, which is difficult to be satisfied by the current cloud rendering architecture. With Asynchronous Timewrp (ATW) and other technologies, the MTP delay requirement could be relaxed to about 70ms. According to the field trial test, for strong interactive VR application when the average MTP delay is less than 80ms, for weak interactive games when the average MTP delay is less than 94ms, the user experience could be basically satisfied. According to the segmentation delay analysis, the service platform rendering and coding delay, network transmission delay can be further optimized space.

2.2.2. Data Rate

The average network transmission data rate required by XR services is generally slightly higher than the service code rate, which is also confirmed by the actual test. In addition, the test found that the service arrival presents Burst characteristics, and the demand for the instantaneous data rate is high. Considering the delay constraint and fluctuation of frame size, he downlink guaranteed bandwidth of the network under 10 ms frame delay constraint is about 2 ~ 3 times of the service code rate

2.2.3. Capacity

In 3GPP simulation, the maximum number of VR users that can be guaranteed in a 100MHz single cell could be more than 10 [5] according to the calculation of service code rate and delay constraint. For VR video, 10 to 12 concurrent users could be supported per cell (100MHz). For strong interactive VR game, the number of concurrent users supported by single cell is less than the theoretical expectation.

2.2.4. Mobility

XR services are highly sensitive to handover delay. The field test results show that the MTP delay is increased at the handover point of 5G network, leading to the phenomenon of screen failure and stutter..

2.3. The Main Problems and Challenges

2.3.1. How to coordinate the network and service

At present, network is not aware the service type or service characteristics of the multimodal data flows generated by XR services, nor can it perceive the importance difference of different data flows at the service layer, so it only provides a general session bearer connection similar to a pipeline. For example, XR downlink audio and video streams are carried by default QoS flow together with other general types of services, and differentiated service guarantee cannot be used.

2.3.2. How to enhance the connection

Based on the performance requirements of XR services measured in Section 2.2, the connection for high-bandwidth and low-latency services could be enhanced from the following aspects.

- As the instantaneous frame-level rate is much higher than the average data rate due to the constraints of frame delay and the frame-peak-to-average ratio, it is necessary to optimize the scheduling according to such service characteristics
- The test results show that network transmission, cloud platform rendering and coding delay account for a large proportion of MTP delay, and there is room for further optimization
- Capacity improvement schemes need to be studied from the aspects of multi-stream collaboration and integrated transmission
- Handover affects XR service experience which may lead to splintered screen, black edge, and stuck, etc, It needs to be optimized from the aspects of reducing handover delay and triggering handover based on frame boundary

2.3.3. How to evaluate the user experience

There is mature quality evaluation solution for voice and video services, but the industry has not yet developed an objective and quantitative evaluation method for the user experience of the XR service. Moreover, different headsets, cloud platforms and network manufacturers have different statistical methods of indicators, which make it difficult to compare horizontally. Therefore, it is urgent to establish a evaluation standard and indicator system, define the mapping model between communication network indicators and user experience indicators, and unify the indicator definition, data collection, measurement and reporting and statistical methods.

3. The XR technology of network, service, device co-operation

3.1. Service Awareness

3.1.1. Feature-based service recognition

When Group of picture (GoP) coding or Slice-based coding is used for XR services, the output stream can be divided into I stream or P stream. The I stream contains the I frame encoded by GOP or by Slice-based, while the P stream contains the P frame encoded by GOP or B frame and the P frame encoded by Slice-based.

The GOP image group contains one I frame and multiple P and B frames, and the Slice-base encoding contains one I frame and multiple P frames. I-stream belongs to the basic stream, which is encoded and decoded independently, The image content of I-stream can be decoded without relying on other image streams. P-stream is a predictive stream or an enhanced stream, and the encoding of the stream depends on the before and after image streams in the I-stream or P-stream. The data size of I-stream is larger than the P-stream because it is encoded separately. Usually, the single-frame I-stream is more than 1.5 times of the P-stream. Using this feature, the I-stream and the P-stream can be identified. After identifying the I-stream and P-stream, the period, size, frame header and end frame of I-frame and P-frame can be further counted.

This service model is usually applicable to XR single service flow mode, I stream and P stream establish the same QoS Flow. There are challenges of accuracy in this way of recognition, and image coding needs to adapt to the above model.

3.1.2. Packet based service recognition

Packet pattern analysis can obtain service information by analyzing the RTP message in the payload of GTP-U. Based on the type of message in the RTP message, I stream and P stream can be determined. The unpacking mode can be combined with packet pattern recognition. On the basis of identifying the I and P streams, the RTP packets in the I streams are further analyzed to confirm that the I frame identification is correct, thereby accurately identifying the frame type and avoiding parsing too many RTP packets. When RTP packets are encrypted, this scheme is not applicable.

In addition, DPI technology is a common method for service detection which detect service feature information based on application layer. When the IP packets, TCP or UDP data stream passes through the core network gateway device, the network device analyzes the access traffic data through DPI, and extracts the data service model, service traffic and other parameters, such as data Host or URL, IP address and port, TCP or UDP, packet length, upstream and downstream, DNS information, etc., and then intelligently analyze the types of services in network interaction.

3.1.3. QoS flow based service recognition

QoS Flow is the smallest granularity of QoS, and different QoS flows provide different levels of service guarantee capabilities. XR services usually contain multiple associated data flows, each with different QoS requirements. When establishing XR service QoS flow, it is necessary to bind 5QI that meets the requirements according to QoS characteristics. The base station can identify service types according to 5QI carried by QoS flow. The 3GPP R17 defines the default parameters of 5QI and QoS for bearable XR services as follow.

Table 1 5QI parameters for XR services (extracted from TS 23.501 Table 5.7.4-1)

5QI Value	Resource Type	Default Priority Level	Packet Delay Budget	Packet Error Rate	Default Maximum Data Burst Volume	Default Averaging Window	Example Services
80	Non-GBR	68	10 ms	10-6	N/A	N/A	Low Delay eMBB applications AR
87	Delay-critical GBR	25	5 ms	10^{-3}	500 bytes	2000 ms	Interactive Service - Motion tracking data,
88		25	10 ms	10^{-3}	1125 bytes	2000 ms	Interactive Service - Motion tracking data
89		25	15 ms	10^{-4}	17000 bytes	2000 ms	Visual content for cloud/ edge/ split rendering
90		25	20 ms	10^{-4}	63000 bytes	2000 ms	Visual content for cloud/ edge/ split rendering

3.1.4. Network slice based service recognition

Terminal identify services based on the traffic descriptor in the URSP, and use corresponding network slice indicators and other parameters to initiate session establishment requests to the network. Based on the network slice indicators provided by the terminal, the network provides connection configuration for the PDU session. This is a 3GPP standardized scheme which could provide accurate recognition and perception of service, and the granularity of service recognition is diverse and flexible.

3.2. Network Adaptation

The wireless environment and network load of 5G networks are constantly changing, and the actual bandwidth supported by wireless channels is not static. In this case, the service data rate generated by the application layer (such as the encoding rate of the media encoder) needs to dynamically adapt to the actual bandwidth supported by the network, so as to ensure that the service delay does not increase unexpectedly.

The congestion feedback mechanism of traditional adaptive encoding is only carried out between the client and server in the application layer. The application client (as the receiver of service data) detects occurrence or relief of congestion , and feedbacks the congestion status information to the application server (as the sender of service data), and the application server adjusts the coding rate after receiving the information. This mechanism, which only detects and feedbacks congestion at application layer, cannot obtain congestion information from intermediate nodes on the communication link, resulting in untimely feedback of congestion information and affecting the guarantee of application layer delay.

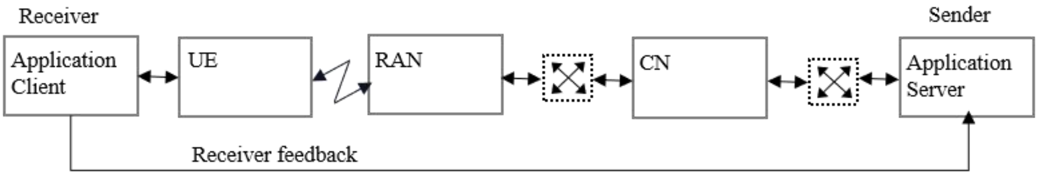

FIG.4 Traditional congestion feedback mechanism

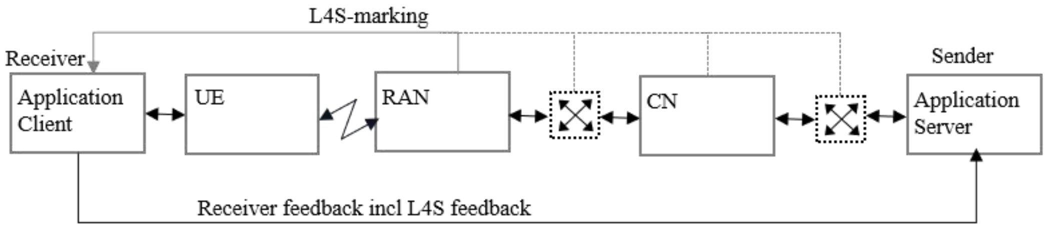

FIG.5 L4S congestion feedback mechanism assisted by 5G RAN

5G-network-based L4S technology combines IETF's L4S (Low-Delay, Low-Loss, Scalable throughput) traffic control scheme with 5G network. 5G base stations monitor network congestion and wireless conditions, and provide congestion information to the application layer by marking the "ECN bit" in the IP packet header, triggering the application layer to dynamically adjust the encoding rate. Compared with the traditional congestion feedback mechanism, the L4S rate adaptation assisted by 5G networks enables the application layer to obtain the congestion status in time, and can assist the application layer to adjust the coding rate in a more timely and efficiently way, thereby avoiding or reduce the burst delay and ensure the user experience.

3.3. Connection Enhancement

This chapter discusses the connectivity enhancement scheme from the capacity enhancement, deterministic delay, mobility enhancement, and power saving.

3.3.1. Capacity Enhancement

(1) Multi-stream collaboration

The data flow of XR and other real-time multimedia services usually have unequal importance. For example, the layered coding in the video compression scheme of H.26x series, encoder encodes data into basic layer with small amount and high importance, and enhanced layer with large amount and low importance. The receiver can restore the video acceptable to the user experience according to the complete basic layer data. On the basis of fully receiving the basic layer data, if it can also receive the enhanced layer data, the user experience can be further improved. Therefore, the network needs to separate transmission and configure different QoS requirements for different importance data flows.

The differentiated transmission based on network-service collaboration can effectively reduce the amount of data required for high reliability, thus reducing the demand for bandwidth and ensuring the basic users experience. At the same time, in the process of network transmission, differentiated transmission guarantees are provided for different transmission requirements of the basic layer and enhancement layer to ensure the synchronization performance between different streams, which can further improve the network transmission efficiency and user experience.

(2) Integrated-frame transmission

In network transmission, each XR service frame will be split into multiple packets. The transmission characteristics of video is that a video frame contains multiple packets, if a data packet is transmitted incorrectly, the video frame may not be decoded correctly. Therefore, when transmitting data packets for XR services, efforts should be made to ensure the overall successful transmission of data packets belonging to the same frame. On the other hand, data in the same task also has integrity requirements. For example, in the haptic Internet, when a robot arm is remotely operated to play basketball, there is dependency between multiple data packets corresponding to video, audio, motion, touch, smell and other information during transmission, and the integrity of task data needs to be guaranteed.

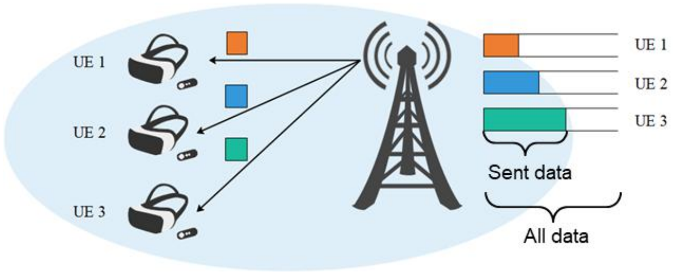

FIG.6 Schematic diagram of proportional scheduling of transmitted data

The 3GPP standard defines an architecture based on network-service collaboration, which carries frame information and key frames through GTP-U. GTP-U contains PDU Set QoS information, including frame number, frame size and end-frame packet identification of XR service, which helps to realize frame integrity transmission on the RAN.

3.3.2. Latency Optimizaiton

(1) Flexible deployment reduces transmission delay

In traditional networks, RAN delay usually occupies only a small part of end-to-end delay, and a large amount of delay is distributed outside of RAN, such as the delay from core network to the service server. For scenarios such as sports venues and theaters, when the high demand for end-to-end delay and certain reliability is required , solutions such as MEC can be used to localize service data, and local data shunting at base station can be adopted to reduce end-to-end delay.

(2) Scheduling optimization reduces RAN delay

3GPP R15, R16 and R17 have introduced delay reduction mechanisms related to URLLC, but these mechanisms are not suitable for high-throughput XR services. Research on scheduling mechanism and efficient resource allocation mechanism suitable for XR service characteristics is the key to reduce delay and increase capacity.

- **Intelligent pre-scheduling**

Based on the learning of the periodic service model, the network tries to match the data transmission rules at UE side. For specific 5QI services that need to be guaranteed, the pre-scheduling can reduce the waiting time for data to be sent, while improving the utilization efficiency of bandwidth. For example, the base station can obtain the service model of uplink service through machine learning, and initiate pre-scheduling within a certain period of time .

- **Configured grant**

For XR each packet size is different and the XR transmission cycle is not matched with the CG cycle, uplink Configured Grant (CG) enhancement is designed in 3GPP R18 to support multiple PUSCH transmission timing and dynamic cancellation mechanism within a single CG cycle. By pre-configuring multiple PUSCH transmission opportunities in the same CG PUSCH, the problem of XR packet size difference and the mismatch between XR transmission cycle and CG cycle can be solved. Less PUSCH transmission resources are used when the packet is small, and more PUSCH transmission resources are used when the packet is large. However, this also faces the problem that PUSCH resources allocated in smaller packets may be wasted, and the dynamic cancellation mechanism is to solve such problems. UE can inform the base station which PUSCH resources are not used through UCI, so the base station can reallocate these unused PUSCH resources to other UEs to improve the efficiency of spectrum utilization.

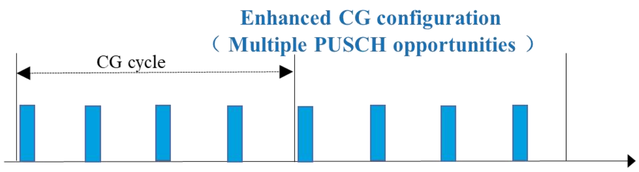

Figure 7 CG enhancement schematic diagram

- **BSR enhancement**

BSR enhancement is to improve the accuracy of BSR to reduce quantization errors, and also to report dealy information of uplink cache data. When the BSR table defined in the 3GPP has large packet sizes and quantization errors, the base station cannot accurately estimate the amount of data to be transmitted on the terminal based on the BSR , resulting in inaccurate allocation of uplink resources. The primary goal of BSR enhancement is to add new BSR entries to reduce the quantization error reported by the BSR.

Secondly, reporting the delay information of uplink cache data can also help XR service to meet its strict packet delay budget. For example, when the UE reports the size of cached data, it can report the remaining packet delay information of the data at the same time, so that the base station can rationally allocate uplink resources according to the remaining delay budget of different users.

(3) Frame-level scheduling provides the low latency

Based on accurate service identification, the user scheduling is enhanced from "packet" scheduling to "frame" scheduling. In addition to considering the priority of PDU Set, it is necessary to schedule between traffic flows of different priority levels on the RAN based on the frame delay requirement.

In downlink scheduling, the transmission resources of different users are determined according to the estimation of RAN capacity and the size of the frames to be used, combined with the remaining margin of each QoS Flow frame delay. The smaller the remaining frame delay margin, the more resources are allocated to maximize the gain and allow more QoS Flow to meet the frame delay requirements.

3.3.3. Mobility Improvement

When XR services gradually move from indoor to outdoor scenarios, such as vehicle XR, it is necessary to study how to meet the strict delay budget requirements in mobility scenarios.

The average frame arrival period of XR services at 60fps is 16.67ms. In the traditional cell handover based on L3, the interrupt delay reaches about 50~60ms. For XR services at 60fps, the interrupt time will cause 3~4 consecutive frames to fail, affecting the user experience. For the future 90fps and 120fps XR services, the interruption caused by the cell handover will lead to more frame failures, resulting in a significant decline in the experience at the handover moment.

(1) Precise control before switching

In the cell handover scenario, it is necessary to use delay scheduling and other methods to ensure the transmission of frames in order to avoid frame delay failed caused by handover interruption. The premise of handover control is that the base station can recognize the frame header and end-frame, and trigger the handover when the previous frame has been received and the next frame that has not yet sent, so as to ensure zero interruption in the frame transmission process.

(2) Agile interaction during handover

L1/L2 switching technology is a cell handover technology that is being standardized by 3GPP Rel-18,to achieve very low interruption delay during handover, and quickly restore the high throughput after handover, which helps to improve the user experience of XR services in mobile scenarios

- The measurement results report and handover decisions between UE and base stations are based on L1/L2 signaling, without submitting to L3 for measurement report decoding, so that the handover process can be completed more timely.
- The downlink synchronization process and uplink synchronization process with the target cell can be completed before handver, realizing the cell handover without random access process and reducing the interruption time.
- The RLC layer can be directly maintained without reconstruction, which reduces the data loss in UM mode and the data retransmittion in AM mode, making L2 data processing more continuous.

3.3.4. Power Saving

The current battery life of XR devices such as VR headsets is about 3 to 4 hours, which is one of the key challenges faced by industry and has a direct impact on the user experience.

From R15 to R18, there are various power-saving mechanisms, in which CDRX enhancement can better match XR services for periodic data transmission (such as audio/video), and reduce terminal power consumption through alternating switching between activity and inactivity. Before 3GPP R17, the parameter value of DRX cycle is an integer multiple millisecond interval between 10ms and 10240ms. There is a problem that the CDRX cycle does not match the XR data packet transmission cycle, resulting in increased data packet delay. The enhanced CDRX mechanism in R18 can support the CDRX cycle of the decimal period, so that the CDRX cycle can be accurately matched to any media frame rate, and power saving gain can be obtained.

3.4. End-Cloud Collaboration

Rendering refers to the process of using software to generate images from models, projecting the models in 3D scenes into digital images according to the set environment, lighting, materials and rendering parameters.. VR applications need to create a completely virtual environment for the user, and the virtual world needs to be rendered at high resolution and frame rate to provide realistic visual effects and smooth animation. Since VR users will be completely immersed in the virtual environment, any delay or lack of smoothness in rendering can cause motion sickness or reduce the quality of the user experience. Therefore, VR has a high demand for rendering capabilities and requires high computing power to render complex graphics and scenes in real time to ensure users get a smooth, authentic feeling. Compared with VR applications, AR overlays virtual content into the real world, users can see enhanced digital information in real time through AR devices, and the rendering of virtual objects needs to be integrated with the real environment, so the complexity of rendering is relatively low. According to the different needs of VR and AR for rendering capabilities, their collaborative rendering schemes are also different.

3.4.1. Collaborative rendering schemes for VR applications

In order to meet the high computing power requirements of VR applications and reduce the terminal computing burden, one method is to implement the End-cloud collaborative rendering scheme based on 5G network. Combining XR device, 5G RAN and edge cloud computing, the real approaching and immersive XR experience is realized through collaborative separation rendering. The computing task is distributed between the edge cloud server and the XR device. On side of the XR device, it mainly performs some energy-saving, low-delay frame rendering and head pose data tracking, while on the side of edge cloud, it focuses on computationally intensive computing processing and graphics rendering. Finally, the results of both sides of the rendering are integrated through 5G networks to provide a truly immersive XR experience.

3.4.2. Separate rendering scheme for AR applications

AR pays more attention to the perception and interactivity of realistic scenes, such as object recognition and spatial positioning, still needs enough rendering capabilities to ensure the seamless integration of virtual content with the real environment and provide a smooth interactive experience. To handle more complex and realistic graphics and scenes, rendering tasks can be carried out in the cloud or edge servers, thereby reducing the power consumption and complexity of the device itself, and providing higher quality rendering effects. At the present stage, the separation rendering of edge cloud-terminal and AR glasses can be realized by connecting AR glasses with 5G terminals and communicating with the cloud. In this architecture, some processing tasks can be offloaded from the AR glasses to the terminal for processing, while more complex and computational-intensive tasks are handled by the edge server. This distributed computing model makes full use of the computing power of the device, and provides users with a smooth and high-quality AR experience with the help of high-performance computing resources of the edge server.

Depending on changes in the quality of 5G wireless channels, rendering tasks can be switched between remote edge cloud computing and local terminal computing. When the 5G signal quality is good, the system can choose to send the computation task to the edge cloud's rendering server for processing to achieve larger-scale computation and resource optimization. And when the 5G signal is weak or the network delay is high, the system transfers the computing task to the terminal for local processing to achieve lower delay. This dynamic switching mechanism based on channel quality enables AR applications to maintain stability under different network conditions, enabling AR glass users to obtain the best computing resource allocation and service experience.

3.5. User Experience Evaluation

For typical XR services, the XR MOS is formed based on user experience evaluation indicators such as experience quality, interaction quality, and source quality, as well as key network KPIs. Combined with quantitative research on experience satisfaction and results of 5G field trial, XR service experience quality classification standards needs to be formed.

3.5.1. User experience indicators

User experience indicators is mainly composed of experience quality, interaction quality and source quality. The experience quality is an indicator that describes the quality of the service presented by XR headset, as well as subjective or objective statistics of user experience when experiencing service, including stall event, screen splashing, black edge and frame loss.

Interactive quality refers to the response quality of the interactive system when the user interacts with the virtual environment. If the interactive system responds quickly and accurately, the user can smoothly explore in the virtual world. If the interactive system responds slowly or inaccurately, users may feel uncomfortable and even suffer from motion sickness. Source quality is to describe the basic performance and objective indicators of the terminal headset and content platform, including the resolution of the headset, refresh rate, volume and weight, content resolution, voice/video coding bitrate and other indicators.

3.5.2. Network Key Performance Indicators

XR services have the characteristics of frame pattern transmission, it is necessary to redefine network key performance indicators for XR immersion service, and unify the statistics and measurement methods of indicators, such as following indicators:

- Frame transmission delay. The time difference between the 5G network receiving the first uplink/downlink data packet of the frame and transmitting the last uplink/downlink data of the frame.
- RAN Frame transmission delay. The time difference between the base station receiving the first uplink/downlink data packet of the frame and transmitting the last uplink/downlink data of the frame.
- Frame data rate. Tthe ratio of the total size of all uplink/downlink data packets received by the cloud/terminal in each frame to the uplink/downlink network transmission frame latency.
- Frame reliability. The ratio of the number of lost frames in a certain period of time to the total number of frames transmitted by the network.

XR user experience indicators are inextricably linked with key network KPI indicators, so it is necessary to study and establish a mapping model between them, For example, the user experiences the stuttering phenomenon, which is mapped to the frame delay and frame data rate indicators of the network, and the screen corruption phenomenon is mapped to the frame reliability of the network, so as to realize the rapid positioning and delimitation of problems.

4. The Value of XR

4.1. Application Scenarios

As a next-generation terminal, XR can support a variety of new applications. Mainstream applications such as watching movies, socializing, and e-commerce will all experience new upgrades due to XR technology.

Scenario 1: Home entertainment. Users can wear XR glasses on the sofa at home or on a long journey to experience the immersive 3D giant screen effect of an IMAX movie theater. The images recorded by users will also be developed from 2D to 3D. XR can better record and play back beautiful moments through 3D images. Through 3D reconstruction volume video and virtual real fusion AR technology, users can project their favorite star image to the real world, such as home and any other scene, fans can take photos with their favorite star virtual real fusion video.

Scenario 2: Online education. The epidemic has promoted the popularity of online education around the world, but the complex structure and abstract concepts of physics, chemistry, biology and astrophysics that are difficult to express through language and flat pictures are more intuitive and understandable through XR terminal combined with three-dimensional models and visualization technology. In addition to attending lectures, hands-on experiments are also a necessary part for learning and understanding. Through three-dimensional interactive technology, students can operate complex physical experiments and understand physical laws.

Scenario 3: Immersive social interaction. Users socialize in the 2D age by sharing text, pictures and videos over the Internet. With the popularity of XR terminal, users can immerse themselves in different 3D virtual spaces (coffee shops, cinemas) to drink coffee and watch movies together . As shown in figure 8, it is XR bringing an unprecedented new experience in 3D space. In the virtual world, users can also have their own three-dimensional avatars. Through 3D scanning and reconstruction technology, users can create a 3D avatar similar to their own image, combined with digital clothing, to form a unique online identity. Through virtual image technology, online virtual socializing eliminates the unfamiliarity of offline socializing, forming a scene of new online socializing.

Scenario 4: Online shopping. The application of XR three-dimensional technology can bring a 360 ° real display of products, combined with AR's 3D interaction, and can also provide consumers with a virtual and real fusion of fitting, comprehensively improving their shopping experience. At present, e-commerce companies have launched new features such as 3D immersive scene shopping, AR fitting, and 3D product display, allowing consumers to grasp 360 ° comprehensive product information, interact with 3D products in real-time, and experience the immersive feeling of the product being "in front of them". The MR terminal combined with the real scene can construct a digital room, virtual "remove" the existing furniture in the room, put 3D furniture models in the real room, and easily complete the virtual and real integration experience of indoor furniture decoration and placement.

4.2. Discussion of XR bussiness model

According to the service types provided by operators, XR scenarios can be divided into industry 2B scenarios, 2B2C mixed scenarios, and individual 2C scenarios. The corresponding bussiness models are discussed below for these three scenarios.

4.2.1. For 2B scenario

2B scenario refers to that services is provided only to industry customers and do not involve industry customers. Its core service model may includes two types, network service and platform service. Network service is to customize network support for customers' needs. The platform service is that telecom operators provide a service platform for XR content providers, and the service platform provides resources and capabilities such as computing power, rendering and technical tools.

The network service fee model could includes charging a fixed monthly fee for the network service, or charging a monthly fee for the network service based on the total duration. In the platform service fee model, the monthly platform capacity rental fee could be charged according to the platform resources and capabilities.

4.2.2. For 2B2C mixed scenario

2B2C mixed scenario refers to providing targeted traffic and network support to individual customers specified by XR industry customers to ensure the experience of target customers. The guaranteed network service fee may be charged according to the user's occupation of network resources, e.g. the number of individual customers or the total duration of network usage.

4.2.3. For 2C scenarios

The individual 2C scenarios refers to that operators directly provide network services for individual customers using XR, or directly provide XR application services to individual users. XR network guarantee service refers to the provision of bandwidth, delay, and other security services by operators for individual customers when using XR applications. XR application service means that operators provide XR application content service directly to individual consumers.

The charging mode of providing network guarantee service is that operators provide network guarantee service packages to meet the short-term or flexible use needs of consumers, and the network guarantee service packages may be charged according to the duration, which can include directional traffic charges. The charging mode of XR application service may take time as the basic charging unit, and provide superimposed rights and interests

5. Summary

This white paper analyzes the requirements of XR service for 5G network in terms of bandwidth, delay and mobility, and discusses key technical solutions including service awareness , network adaptation , connection enhancement, End-Cloud collaboration and user experience evaluation. The purpose of this white paper is to reach industry consensus, jointly promote product implementation and technology maturity by industry partners, and ensure the development of XR with the distinguished performance of 5G network.

Looking forward to the future, GTI will continue to carry out collaborative technology innovation, product innovation and application innovation together with industrial partners, so as to accelerate the large-scale application of XR, provide users with immersive experience anytime and anywhere.

References

[1] China Mobile Research Institute, “5G Cloud XR End-to-End Capability Requirements”, 2020

[2] China Mobile Research Institute, “Research Report on 5G Cloud XR Cloud Network Architecture and Solution”, 2020

[3] IETF RFC 9330 "Low Latency, Low Loss, Scalable Throughput (L4S) Internet Service: Architecture".

[4] 3GPP TR 26.928: “Extended Reality (XR) in 5G”

[5] 3GPP TR 38.838: “Study on XR (Extended Reality) Evaluations for NR”

[6] 3GPP TS 23.501: “System architecture for the 5G System (5GS)”

Abbreviations

Abbreviations	Full Name
XR	Extended Reality
VR	Virtual Reality
AR	Augmented Reality
MR	Mixed Reality
AI	Artificial Intelligence
FOV	Field of View
MTP	Motion to Photons
KPI	Key Performance Indicator
GoP	Group of picture
P frame	predictive-frame
B frame	Bi-directional interpolated prediction frame
I frame	intra picture
GTP	GPRS Tunneling Protocol
RTP	Real-time Transport Protocol
PDU	Protocol Data Unit
DPI	Deep Packet Inspection
TCP	Transmission Control Protocol
UDP	User Datagram Protocol

Abbreviations	Full Name
URL	Uniform Resource Locator
DNS	Domain Name Syste
QoS	Quality of Service
5QI	5G QoS Identifier
URSP	UE Route Selection Policy
IETF	The Internet Engineering Task Force
L4S	Low-Latency, Low-Loss, Scalable throughput
ECN	Explicit Congestion Notification
RTT	Round-Trip Time
RAN	Radio Access Network
URLLC	Ultra-Reliable Low-Latency Communications
CG	Configured Grant
PUSCH	Physical Uplink Shared Channel
UCI	Uplink Control Information
UE	User Equipment
BSR	Buffer Status Report
FPS	Frame Per Second
PDCP	Packet Data Convergence Protocol
RLC	Radio Link Control
LTM	L1/L2 Triggered Mobility
UM	Unacknowledged Mode
AM	Acknowledged Mode
DRX	Discontinuous Reception
CDRX	Connected-mode Discontinuous Reception
MOS	Mean Opinion Score
IMAX	Image Maximum
2B2C	To Business To Consumer