True 3D Holography: A Communication Service of Tomorrow and Its Requirements for a New Converged Cloud and Network Architecture on the Path to 6G

Ingo Friese*, Mandy Galkow-Schneider*, Louay Bassbouss†, Alexander Zoubarev†, Andy Neparidze†, Sergiy Melnyk‡, Qiuheng Zhou‡, Hans D. Schotten‡§, Tobias Pfandzelter¶, David Bermbach¶, Arndt Kritzner‖, Enrico Zschau**, Prasenjit Dhara††, Steve Göring††, William Menz††, Alexander Raake††, Wolfgang Rüther-Kindel‡‡, Fabian Quaeck‡‡, Nick Stuckert‡‡, Robert Vilter‡‡,

* Deutsche Telekom AG, Berlin, Germany
Email: {ingo.friese, mandy.galkow-schneider}@telekom.de
† Fraunhofer FOKUS, Berlin, Germany
Email: {louay.bassbouss, alexander.zoubarev, andy.neparidze}@fokus.fraunhofer.de
‡ Intelligent Networks, German Research Center for Artificial Intelligence (DFKI GmbH), Kaiserslautern, Germany
Email: {sergiy.melnyk, qiuheng.zhou, hans_dieter.schotten}@dfki.de
§ Institute for Wireless Communication and Navigation, University of Kaiserslautern-Landau, Kaiserslautern, Germany
Email: schotten@rptu.de
¶ Mobile Cloud Computing Group, Technische Universität Berlin & ECDF, Berlin, Germany
Email: {tp, db}@mcc.tu-berlin.de
‖ Logic Way GmbH, Schwerin, Germany
Email: kritzner@logicway.de
** SeeReal Technologies GmbH, Dresden, Germany
Email: ez@seereal.com
†† Audiovisual Technology Group, Technische Universität Ilmenau, Ilmenau, Germany
Email: {prasenjit.dhara, steve.goering, william.menz, alexander.raake}@tu-ilmenau.de
‡‡ Aviation Engineering Group, Technical University of Applied Sciences Wildau, Wildau, Germany
Email: {wkindel, quaeck, stuckert, vilter}@th-wildau.de

***Abstract*—Research project *6G NeXt* is considering true 3D holography as a use case, setting requirements on the communication as well as the computing infrastructure. In a future holographic communication service, clients are widely spread in the network and cooperatively interact with each other. Especially for holographic communication high processing power is required as well. This makes a high-speed distributed backbone computing infrastructure, which realizes the concept of split computing, inevitable. Furthermore, tight integration between processing facilities and wireless networks is required in order to provide an immersive user experience. This paper illustrates true 3D holographic communication and its requirements. Afterward, an appropriate solution approach is elaborated. Here, novel technological approaches are discussed based on a proposed overall communication and computing architecture.**

***Index Terms*—6G, agile link adaptation, split computing, high-performance backbone, holographic communication, true 3D holography, quality of experience, remote rendering, discovery service, resource broker, AI-supported flexible mobile communication platform**

I. Introduction

The holographic 3D (H3D) use case shows a great example of an immersive virtual technology that underlines the need to converge computing and connectivity. The research project *6G NeXt* [1] develops an integrated modular architecture enabling future applications to find the most suitable resources in the cloud and network. This ability is crucial, especially for an application that needs to split and distribute computing tasks. The proposed architecture involves a set of functionalities, algorithms, and tools for coordinating and managing the execution of these computations across multiple geographically distributed cloud deployments and networks.

The rest of this paper is structured as follows: Section II describes existing concepts and how the proposed architecture relates to them. Section III illustrates Holographic Communication as a use case based on remote rendering as an example of split computing. Section IV proposes an architecture that supports applications and services to distribute computing tasks across client, edge, and central cloud deployments, taking optimal connectivity into account. This includes the following parts of the proposed architecture:

A) A remote rendering framework as an example of split computing
B) A combined high-performance computing and communication backbone helping clients and services to find the most sufficient resources in cloud and network
C) Edge-to-cloud services and software extensions to support the architecture
D) A Quality of Experience (QoE) alignment ensures a great user experience describing the necessary precondition all over the system
E) An AI-supported flexible mobile communication platform

Section V concludes this paper.

II. State of the Art and Related Work

The idea to split computing tasks of eXtended Reality (XR) applications is not new. Many Virtual Reality (VR) and Augmented Reality (AR) devices do not have sufficient CPU/GPU power and memory to render complex virtual scenes in real time. *3GPP* standardized in release 18 of its 5G specification an *Architecture for enabling Edge Applications* [2]. The goal of this architecture is to host applications in an edge cloud close to the base station and thus near to the clients in order to reduce end-to-end latency. The proposed architecture of this paper complements the *3GPP* standard towards convergence of computing and connectivity across the whole cloud continuum.

III. HoloCom use case

The *6G NeXt* project is going to implement a real prototype of a future holographic communication service. The idea and goal of the use case called *Holographic Communication* (*HoloCom*) is to research and develop a 3D video communication system based on true H3D displays, as shown in Figure 1. We aim for realistic and natural remote communication between individuals enabled by H3D technology invented by one of our partners. This technology creates light information in space in the same way as real objects would do, to allow natural viewing for the human eye including continuous depth, eye-focus, and convergence. This enables a natural audiovisual communication experience by reconstructing the call counterpart as a live and true 3D hologram in 3D space. Integrated eye tracking allows free movement and changing of perspective with little to no noticeable delay. Instead of using a H3D display VR headsets could be optionally used, although this limits options to capture the person's head and face for presentation on the other side.

Figure 1. Holographic Communication

The *HoloCom* use case requires 3D capturing of real persons in real-time, performed on each of the participating client systems. The 3D data will be transformed into a live 3D model to be synthesized into a virtual environment within a 3D engine running on the edge domain realized as a headless *Unity* 3D rendering service [3].

A *Unity 3D specific plugin* to capture configurable 3D views compatible for H3D is part of the development. The generated data is finally transferred to the other client using real-time streaming technologies such as WebRTC [4] or RTSP [5]. To enable this type of transfer, the captured H3D compatible views are encoded into an image-based format. On the client side, the incoming 3D data is then computed into holograms and presented on the H3D display or alternatively on a VR headset.

The generated H3D views depend on the real eye location of the observer looking at the H3D display. Thus, with changing eye location, the H3D views and accordingly the holograms need to be updated with high frequency. This underlines the challenging low delay and high bandwidth requirements that we expect in the *HoloCom* use case. Extremely low delay combined with high bandwidth enables a real-time eye movement to H3D display photon latency enabling a realistic experience highlighting the requirement to 6G technology. Higher bandwidth clearly allows for faster transfer time to reduce delay and higher quality 3D data to be transferred to increase overall experience and quality of the person representation/reconstruction in full 3D.

Some key research will also be done in the field of 3D capturing with several cameras of different types. Normal RGB sensors to *time-of-flight* or *depth-from-stereo* based depth capturing methods will be considered in the research with the result of generating a live, high-resolution 3D model of a person. This will be based on the technology of one of our partners, which already enables high-quality 360 degree live scanning, producing more than 10 terabytes of data per minute. Key challenges will be clean capturing of, e. g., face details, hair, or glasses. Finally, audio capturing and hologram playout synchronization will be challenges to be solved.

IV. Concept of a Multi-level Split Computing Architecture

In this section, we outline key concepts of our proposed *HoloCom* architecture.

A. Remote Rendering

Remote rendering is an approach for migrating graphical computation tasks from consumer devices to remote servers, allowing even complex and computation-intensive 3D applications to be experienced on devices with limited processing capabilities, such as those of TVs. These devices typically have the capability to play high-resolution videos up to 4K or 8K, but are not designed for running applications with complex 3D rendering, which requires GPU acceleration.

In remote rendering, the viewport of the virtual camera within the 3D experience is rendered headless on the server and encoded as a video/audio stream, which is then delivered to the user device, which only needs to play the stream. User interactions, which can vary between different device classes, such as mouse/keyboard input on desktops, touch inputs on mobile devices, remote control on TVs, and motion control on AR/VR displays, must be captured and sent to the remote rendering server, which triggers the received events on the underlying rendering engine as if they were received from a connected input device.

The main challenge of remote rendering is reducing the end-to-end (E2E) latency while maintaining image quality. This latency, also known as *Motion-to-Photon (MTP)* or *Click-to-Photon (CTP)* latency, represents the delay until a user interaction is reflected on the display. Many factors affect the end-to-end latency and the quality of a remote rendering session under various conditions and constraints, which are summarized below:

- Network: WiFi, Ethernet, LTE, 5G
- Browser: All modern browsers supporting WebRTC (Chrome, Firefox, Safari, Edge)
- Platform: Mobile, Desktop, TV, XR/VR, Holographic display
- Deployment: On-premise, Edge, Cloud or any combination of them
- Video codecs: VP8, VP9, AV1, H264, H265
- Audio codecs: Opus, AAC
- Encoding/Decoding: Hardware Encoding/Decoding, Software Encoding/Decoding
- Server Resources (GPU/CPU/Memory): Various combination of accelerated compute instances provided by Hyperscalers
- Graphics Engine: Unity, Unreal Engine, NVIDIA Omniverse
- Protocol: WebRTC, HTTP3/QUIC

Figure 2 shows the overall architecture of remote rendering. In order to reduce the end-to-end latency mentioned above, it is important to run the rendering instances close to the consumer device. This is facilitated by the edge computing capability of 5G networks and optimized in future 6G networks by supporting optimized and dynamic deployment of computation tasks across various cloud deployments.

B. High-Performance Communication and Computing Backbone

The performance and efficiency of a new network generation will be determined by high-performance radio interfaces with application-optimized radio protocols as well as by ultra-fast software stacks, intelligent media processing, and the deep integration of artificial intelligence (AI) to optimize the overall system. In the following, the key technologies for a joint communication and computing infrastructure are presented.

1) Split Computing: In the *HoloCom*s use-case, capturing and rendering processes are distributed between a client device, edge, and central cloud deployments. Distribution of computation tasks in the backbone layer

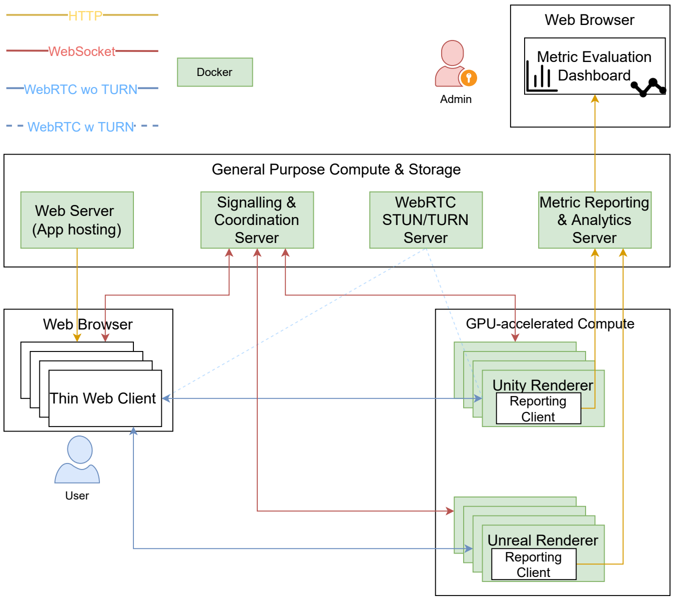

Figure 2. Remote Rendering Architecture

and network based on key performance indicators (KPIs) enhances the concept of edge computing and provides new possibilities for future application. The core of this concept is to match the requirements of an application such as *HoloCom* with the most suitable resources that are available based on KPIs.

2) Cross-Layer Metrics Function: Before a computing task such as rendering can be assigned to a certain computing infrastructure, it needs distinct knowledge about available resources and free capacities. A new proposed metric function aims at gathering metrics and measurement data from three different layers of the overall system, as shown in Figure 3:

a) Application Layer: One part of the metric function collects data that are specific for *HoloCom* application. Typical data for the capturing part of the application include but are not limited to video texture resolution, mesh polygons number, mesh bit rate, and audio bit rate. Rendering-specific data encompasses, e. g., the number of virtual cameras in the scene, used graphics engine, rendering frame rate, or average video decoding time.

b) Computing Layer: This part of the metric function collects performance data of the computing infrastructure such as available CPU or GPU in the client device but also in the distributed edge-to-cloud environments.

c) Network Layer: The metric function also measures typical network characteristics such as throughput, latency, jitter, and packet loss between *HoloCom* clients and various instances of the rendering framework.

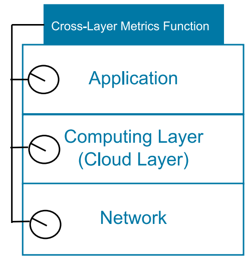

Figure 3. Cross-Layer Metrics Function

All data gathered by the cross-layer metric function are used for performance management of the *HoloCom* application but also for the entire computing and communication backbone and the optimization of the overall system.

3) Discovery Service Function: A *Discovery Service* [6] is a typical part of a microservice architecture. Services such as the *HoloCom* rendering framework might change a number of instances and locations in a virtual and containerized environment over time. The *Discovery Service* provides *HoloCom* clients with the network address of a remote rendering service. Therefore, every instance of a remote rendering service has to register itself with the *Discovery Service* by sending its current address. This is typically the dynamically assigned network path.

The proposed architecture takes the concept of *Discovery* one step further and enables a client to send parameterized *Discovery Request*. A client might request an instance of a rendering service that fulfills certain KPIs. These KPIs might be regarded to bandwidth or latency requirements but also to quality of service (QoS) classes or service costs. The proposed discovery service function keeps track of the instance addresses but takes also data of the *Cross-layer Metric Function* into account in order to provide *HoloCom* clients with the most suitable rendering framework instance in order to assure an optimal service experience.

4) Service Broker Function: A service broker helps services to find, e. g., free cloud capacities and also to start a new instance based on requirements described with KPIs. It is an enhanced version of a discovery service. When the remote rendering framework recognizes a decreasing service quality due to missing memory and CPU/GPU in the current cloud environment it might ask the Service Broker Function for a new deployment possibility. The Service Broker Function performs two tasks: First, it recommends a suitable cloud resource with sufficient connectivity for increasing service capacities.

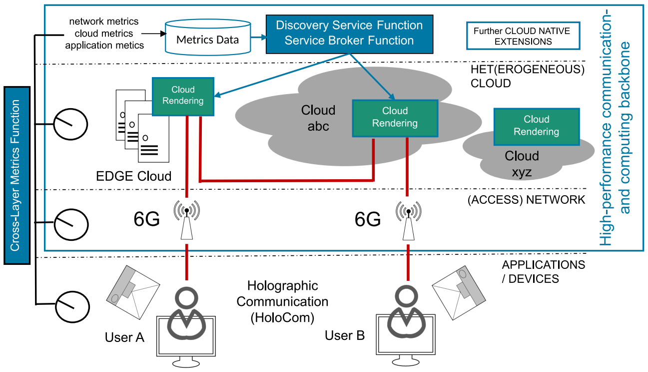

Figure 4. High-performance communication-and-computing backbone for holographic communication

Second, it helps the service to deploy and start a new service instance using Docker [7] and automate infrastructure tools. Also, the network connectivity might be configured using slicing capabilities or network APIs, e. g., those specified in CAMARA [8].

5) Comprehensive Architecture View: The *HoloCom* use-case running on the proposed architecture is shown in Figure 4:

The *HoloCom* user clients contact the *Discovery Service Function* providing their minimal requirements regarding latency, possibly caused by the network and remote rendering service. The *Discovery Service Function* provides user client A with the address of a remote rendering service located in the edge-cloud while client B is connected to a remote rendering service in central cloud deployment abc. The *Discovery Service Function* provides its recommendation based on metrics data gathered by the *Cross-Layer Metrics Function*. The remote rendering service starts a new instance in cloud xyz with the help of the *Service Broker Function*. The architecture is characterized by three new functions added, the *Cross-Layer Metrics Function*, the *Discovery Service Function*, and the *Service Broker Function* as extensions to standard cloud environments. Further extensions are going to be described in the next section.

The classic separation of network and computing infrastructure in layers is eliminated and replaced by a modular approach. The clear distinction between computing infrastructures also starts to blur. The *High-performance communication-and-computing backbone* enables a flexible split of computing tasks according to the requirements of the application. And applications get all information and support to decide where to run, where to connect or to be deployed to.

C. Edge-to-Cloud Serverless Software Extensions

The proposed system leverages existing serverless software platform abstractions for compute, data, and message management in the 6G edge-to-cloud continuum to handle irregular and supplementary tasks such as metadata management.

1) Compute: In addition to the resource-intensive tasks of cloud holographic rendering, the metrics functions that provide ancillary system measurement data must be deployed in a scalable, isolated, and resource-efficient manner throughout the edge-to-cloud continuum. Specifically, metrics must be collected at the client side, within the network, and at the remote rendering location. To deploy this irregular task with both geographical scalability and isolation between tenants, we rely on the Function-as-a-Service (FaaS) compute paradigm, where small, stateless functions can be deployed in an event-driven manner. Using the *tinyFaaS* [9] lightweight edge FaaS platform, we can efficiently provide a software infrastructure to collect critical performance metrics data. Additionally, aggregation and filtering of metrics data can be performed as FaaS functions in the edge-to-cloud continuum to reduce the amount of data that must be transmitted to the service broker.

2) Messaging: After collecting metrics data in geographically distributed locations, it must be distributed efficiently to where it is needed for decision-making. To address this challenge, we leverage the Publish/Subscribe (Pub/Sub) paradigm where services and application components can subscribe to relevant metrics data that individual metrics functions publish. Crucially, through the introduction of geographical awareness and intelligent message routing as well as the early aggregation and filtering of metrics data, we reduce the overall network strain of distributing metrics data. For these tasks, we rely on the *GeoBroker* [10], [11] Pub/Sub middleware.

3) Data: While the holographic data that is streamed through the system may be ephemeral, there is supporting data and metadata that is crucial to the operation of the system. This data includes client authentication information, as well as other critical data that must be accessed in a low-latency manner. Furthermore, clients and services may move fluidly throughout the edge cloud and the real world, and data copies must move with them. In order to replicate such metadata at the client side and the rendering service, and in order to do so even with service or client movement, we rely on the abstractions provided by the *FReD* [12]–[14] fog data management system. Using application-controlled data replica placement, we can ensure low-latency metadata access without impairing user experience, adhering to QoS and QoE objectives.

D. QoE Alignment

The user experience is a central factor for the subsequent satisfaction of users in immersive communication. In the context of *HoloCom*, humans may be sensitive to errors in the displayed information or more subtle impairments due to delay. Along with this, higher values of QoS do not necessarily ensure higher QoE [15]. Hence, it is necessary to conduct detailed user studies on the perceived quality of *HoloCom*. The perceived quality of immersive communication is to be evaluated through subjective as well as objective measures. The subjective measures will be assessed using questionnaires related to constructs such as co-presence, immersion, and realism, as well as in terms of audiovisual quality, which is often measured with a 5-point Absolute Category Rating resulting in a mean opinion score (MOS) [16]. Along with this, we will record conversational and behavioral data of the participants for indirect evaluation of user experience [17]. The QoE of users is related with underlying characteristics of the end device and the network and edge technology used (QoS) to identify respective KPIs.

A joint architecture for communication and computation is proposed for *HoloCom* systems, where the computation of the data from capturing to rendering is distributed across cloud deployments and networks. The parameters of transmission (e. g., codecs, compression ratio and bandwidth, delay, and possible data loss) will be collected to develop a prediction model, relating the aforementioned QoE indicators with such service-related KPIs. Existing state-of-the-art monitoring models such as the framework of the ITU-T P.1203 and P.1204 model series will be used as a conceptual basis, adapted to our context [18], [19]. Such QoE models can serve to automatically measure the current QoE of the application and inform about root-causes, enabling AI-assisted communication system to allocate adequate resources for maintaining user satisfaction of the *HoloCom* systems.

At the receiver side, the rendering of the communication partner is distributed across the computing resources in the cloud/edge networks for efficient processing of the data. However, synchronization of the processed data stream from distributed media processing functions is important for rendering a communication partner. If the data streams from the distributed media processing functions are not synchronous, it might lead to misaligned representations of the rendered object [20]. Also, it is highly desirable to have a synchronous audio signal along with the visual representation to support better QoE in the *HoloCom* systems. Along with this, undesirable values of transmission parameters (e. g., higher data loss, longer delay) severely impact the quality of rendered objects and produce possible artifacts such as blurring, freezing, etc., which negatively impact the user experience.

E. AI-Supported Flexible Mobile Communication Platform

Since wireless networks are now rapidly growing with massive network items, wide bandwidth, and high data rates, it is necessary to take the uncertainty and the changing environment of the radio access network into consideration to improve the system reliability, QoS and even QoE. Recently, machine learning (ML) models are increasingly being used to solve those non-linear problems, e.g., network slicing, channel prediction, or QoS optimization. Because of those advantages for adapting the spatiotemporal dynamics in advance based on the pre-trained model on a huge amount of historical channel data [21]–[27]. In communication networks with moving objects, the experienced QoS is affected by various factors, such as User Equipment (UE) density, interference, mobility, handover, and roaming transitions. The prediction of worse network services at certain locations due to one or many of those factors is one of the key requirements for critical applications with high demand of QoS/QoE. A recent approach has combined the most efficient ML methods, that an long short term memory (LSTM) and a deep reinforcement learning (DRL) are used to conduct online decentralized testing at the UE pairings, which gives resource units the ability to allocate bandwidth and make decisions about packet scheduling [28]. Further, federated deep reinforcement learning (FDRL) is proposed for a UAV-aided vehicular network, which is examined to improve connection and minimize latency [29]. ML methods are used at every level in distributed federated systems, which reduces latency and improves accuracy, which are the most important indexes of the quality of *HoloCom*.

An overview of such a ML-based prediction model is shown in Figure 5.

1) Offline phase: In this phase, QoS data are collected using network and application monitoring tools, which are processed to train and periodically update a prediction model. Different data and input features need to be used for an accurate prediction of different QoS parameters. The objective of this phase is to leverage

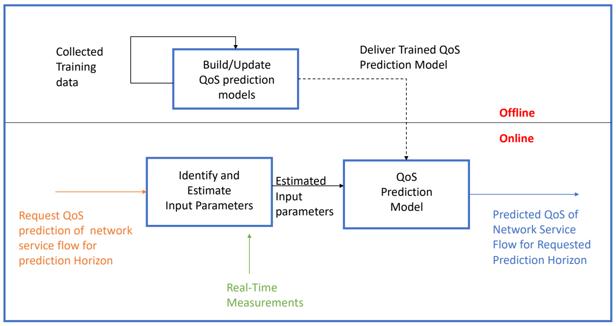

Figure 5. Overview of QoS prediction scheme for network service flow

machine learning techniques to establish connections between input features and their corresponding predicted parameter and estimate the anticipated value of a QoS parameter with greater accuracy. This phase is conducted offline and does not recur for each individual prediction request, although it may be periodically re-executed based on the latest collected data. It may be necessary to develop distinct prediction models for each QoS parameter and/or for Uplink and Downlink signals. The trained models are then made available to the online engine.

2) Online phase: This phase is triggered by each prediction request. Based on the type of request, the relevant input features are identified, and their expected values are projected for the prediction horizon specified by the application. Real-time measurements are also gathered by the appropriate network entities to determine the present and future values of the input features for the given service flow. In the second stage of this phase, the anticipated values of the input features are inputted into the trained QoS prediction model, which calculates the anticipated QoS of a service flow for the requested prediction horizon.

In Figure 4, a user connects to the *remote rendering* service instance by means of a wireless link. However, on the application layer, the communication link may be divided into different traffic types, such as video, audio, control, etc. Every of these traffic types has its dedicated requirements on throughput, reliability, latency, and others. The variety of the traffic types becomes even larger if some other cloud-based applications are active in parallel to a *HoloCom* call. The very heterogeneous requirements of different applications and in particular different traffic types inside a single application require high flexibility of the underlying wireless network in order to provide optimal QoE to the user.

Especially, interactive applications such as *HoloCom* require high link reliability in conjunction with low latency. This means, that the approach of retransmission of lost packets is not appropriate, but the packets should arrive on time with the given loss probability. Besides, the well-established adaptation of Modulation and Coding Scheme (MCS) to the quality of the wireless channel does not provide the required flexibility, since the jump between two MCSs is too high in terms of throughput change.

To overcome this shortage, we propose the utilization of fountain codes, which belong to the family of rateless codes. Here, an arbitrary number of sub-codewords is generated out of a payload packet, but only a certain number of those is required to decode the packet. Thus, the coding rate can be flexibly adjusted. However, these codes do not include forward error correction (FEC) capability thus are not able to detect nor to correct corrupted sub-codewords. In order to overcome this shortage, a two-stage coding approach can be applied. In the first stage, FEC codes are utilized to turn the wireless channel into a packet erasure channel. In the second stage, fountain coding is utilized to correct occurred packet drops. Authors in [30] showed, that the joint erasure and error-correcting scheme shows better performance as compared to the one-stage approach.

The main advantage of fountain codes is their flexibility. Based on predicted channel quality, the coding rate can be adjusted “on the fly” in order to meet the required reliability [31]. Since the coding rate is not related to MCS settings, it can be re-adjusted for any piece of payload. It provides the possibility to divide the heterogeneous traffic types into separate communication queues. After, based on predicted channel quality and the requirements of each queue, the coding scheme can be flexibly adjusted to meet the QoS requirements on the one hand, and to provide optimal utilization of communication channels on the other hand.

V. Conclusion

Sixth generation networks should bring mobile communication to the next level making the connectivity omnipresent for commercial and private usage. In order to realize this vision, new technological ideas are required. In this paper, we presented an approach and an architecture of joint communication and computing infrastructure. The proposed high-performance backbone layer in combination with additional cloud native services allows the distribution of processing power in all levels inside the communication network, which reduces the demands on the client devices on the one hand, and simplifies the realization of collaborative applications on the other hand. With *HoloCom* a demanding use case was chosen to provide real-world requirements on latency, throughput, reliability, availability of computation power etc. The goal of the project *6G NeXt* is to go beyond the conceptual evaluation of proposed technologies towards a functional proof of concept based on the proposed *HoloCom* use case.

ACKNOWLEDGEMENTS

Funded by the Bundesministerium für Bildung und Forschung (BMBF, German Federal Ministry of Education and Research) – 16KISK174K, 16KISK176, 16KISK177, 16KISK178, 16KISK179, 16KISK180, 16KISK182, 16KISK183.

REFERENCES

[1] "6gnext project homepage." [Online]. Available: https://6gnext.de/

[2] 3GPP, "Architecture for enabling Edge Applications," 3rd Generation Partnership Project (3GPP), Tech. Rep. 23.558, Mar. 2023, version 18.2.0.

[3] "Unity homepage." [Online]. Available: https://unity.com/

[4] "Webrtc homepage." [Online]. Available: https://webrtc.org/

[5] IETF, "Real Time Streaming Protocol (RTSP) RFC 2326," Internet Engineering Task Force, Tech. Rep. 2326, 1998, update 2022-04-07.

[6] "Service discovery in microservices," Nov. 2022. [Online]. Available: https://www.baeldung.com/cs/service-discovery-microservices/

[7] "Docker homepage." [Online]. Available: https://www.docker.com/

[8] "Camara project homepage." [Online]. Available: https://camaraproject.org/

[9] T. Pfandzelter and D. Bermbach, "tinyFaaS: A lightweight faas platform for edge environments," in *ICFC 2020*, Apr. 2020, pp. 17–24.

[10] J. Hasenburg and D. Bermbach, "GeoBroker: Leveraging geo-context for IoT data distribution," *Elsevier Computer Communications*, vol. 151, pp. 473–484, 2020.

[11] ——, "DisGB: Using geo-context information for efficient routing in geo-distributed pub/sub systems," in *UCC 2020)*, Dec. 2020.

[12] T. Pfandzelter *et al.*, "Managing data replication and distribution in the fog with fred," *arXiv:2303.05256 [cs.DC]*, 2023.

[13] J. Hasenburg *et al.*, "Fbase: a replication service for data-intensive fog applications," *arXiv:1912.03107 [cs.DC]*, 2019.

[14] ——, "Towards a replication service for data-intensive fog applications," in *ACM SAC*, 2020, pp. 267–270.

[15] M. Agiwal *et al.*, "Next generation 5g wireless networks: A comprehensive survey," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 3, pp. 1617–1655, 2016.

[16] S. Aseeri and V. Interrante, "The influence of avatar representation on interpersonal communication in virtual social environments," *IEEE Transactions on Visualization and Computer Graphics*, vol. 27, no. 5, pp. 2608–2617, 2021.

[17] J. Skowronek *et al.*, "Quality of experience in telemeetings and videoconferencing: a comprehensive survey," *IEEE Access*, 2022.

[18] W. Robitza *et al.*, "Http adaptive streaming qoe estimation with itu-t rec. p. 1203: Open databases and software," in *Proceedings of the 9th ACM Multimedia Systems Conference*, 2018, pp. 466–471.

[19] A. Raake *et al.*, "Multi-model standard for bitstream-, pixel-based and hybrid video quality assessment of uhd/4k: Itu-t p. 1204," *IEEE Access*, vol. 8, pp. 193 020–193 049, 2020.

[20] A. Clemm, M. T. Vega, H. K. Ravuri, T. Wauters, and F. De Turck, "Toward truly immersive holographic-type communication: Challenges and solutions," *IEEE Communications Magazine*, vol. 58, no. 1, pp. 93–99, 2020.

[21] J. Perdomo *et al.*, "Qos prediction-based radio resource management," in *2022 IEEE 96th Vehicular Technology Conference (VTC2022-Fall)*, 2022, pp. 1–6.

[22] Q. Zhou *et al.*, "Performance evaluation over dl-based channel prediction algorithm on realistic csi," in *2022 IEEE 96th Vehicular Technology Conference (VTC2022-Fall)*, 2022, pp. 1–5.

[23] W. Jiang and H. D. Schotten, "Neural network-based fading channel prediction: A comprehensive overview," *IEEE Access*, vol. 7, pp. 118 112–118 124, 2019.

[24] R. Levie *et al.*, "Radiounet: Fast radio map estimation with convolutional neural networks," *IEEE Transactions on Wireless Communications*, vol. 20, no. 6, pp. 4001–4015, 2021.

[25] D. F. Külzer *et al.*, "Cdi maps: Dynamic estimation of the radio environment for predictive resource allocation," in *2021 IEEE 32nd Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, 2021, pp. 892–898.

[26] W. Jiang and H. D. Schotten, "Deep learning for fading channel prediction," *IEEE Open Journal of the Communications Society*, vol. 1, pp. 320–332, 2020.

[27] Q. Zhou *et al.*, "Deep learning-based signal-to-noise ratio prediction for realistic wireless communication," in *2022 IEEE 95th Vehicular Technology Conference: (VTC2022-Spring)*, 2022, pp. 1–5.

[28] M. Khayyat *et al.*, "Advanced deep learning-based computational offloading for multilevel vehicular edge-cloud computing networks," *IEEE Access*, vol. 8, pp. 137 052–137 062, 2020.

[29] H. Alsulami *et al.*, "A federated deep learning empowered resource management method to optimize 5g and 6g quality of services (qos)," *Wireless Communications and Mobile Computing*, vol. 2022, pp. 1–9, 2022.

[30] M. Karrenbauer *et al.*, "A study on the application of rateless coding in non-cellular mimo systems for machine-type communication," *IFAC*, vol. 53, no. 2, pp. 8243–8248, 2020.

[31] S. Melnyk *et al.*, "Wireless industrial communication and control system: Ai assisted blind spot detection-and-avoidance for agvs," in *IN4PL*, Mar. 2022, pp. 1301–1306.