http://dx.doi.org/10.25673/111638

6G NeXt — Joint Communication and Compute Mobile Network: Use Cases and Architecture

Sergiy Melnyk,[1] Qiuheng Zhou,[1] Hans D. Schotten[1,2] Wolfgang Rüther-Kindel,[3] Fabian Quaeck,[3] Nick Stuckert,[3] Robert Vilter,[3] Lisa Gebauer[3], Mandy Galkow-Schneider,[4] Ingo Friese,[4] Steffen Drüsedow[4], Tobias Pfandzelter,[5] Mohammadreza Malekabbasi,[5] David Bermbach[5], Louay Bassbouss,[6] Alexander Zoubarev,[6] Andy Neparidze[6], Arndt Kritzner[7], Jakob Hartbrich,[8] Alexander Raake[8], Enrico Zschau[9], Klaus-Jürgen Schwahn[10]

Abstract:

The research on the new generation mobile networks is currently in the phase of defining the key technologies to make 6G successful. Hereby, the research project *6G NeXt* is aiming to provide a tight integration between the communication network, consisting of the radio access as well as backbone network, and processing facilities. By the concept of split computing, the processing facilities are distributed over the entire backbone network, from centralised cloud to the edge cloud at a base station. Based on two demanding use cases, *Smart Drones* and *Holographic Communication*, we investigate a joint communication and compute architecture that will make the application of tomorrow become reality.

Keywords: 6G, agile link adaptation, split computing, high-speed backbone, geo-distributed computing, uav, anti-collision system, wireless closed loop, holographic communication, quality of experience, split rendering, metaverse

1 Introduction

The idea of offloading demanding processing tasks to powerful machines by splitting an application into client and server parts has been known for decades. Nevertheless, the

---

[1] German Research Center for Artificial Intelligence (DFKI GmbH), Kaiserslautern, Germany, {sergiy.melnyk, qiuheng.zhou, hans_dieter.schotten}@dfki.de

[2] University of Kaiserslautern-Landau, Kaiserslautern, Germany, schotten@rptu.de

[3] Technical University of Applied Sciences Wildau, Wildau, Germany, {wkindel, quaeck, stuckert, vilter, lisa_josephine.gebauer}@th-wildau.de

[4] Deutsche Telekom AG, Berlin, Germany, {mandy.galkow-schneider, ingo.friese, steffen.druesedow}@telekom.de

[5] Technische Universität Berlin & ECDF, Berlin, Germany, {tp, mm, db}@mcc.tu-berlin.de

[6] Fraunhofer FOKUS, Berlin, Germany, {louay.bassbouss, alexander.zoubarev, andy.neparidze}@fokus.fraunhofer.de

[7] Logic Way GmbH, Schwerin, Germany, kritzner@logicway.de

[8] Technische Universität Ilmenau, Ilmenau, Germany, {jakob.hartbrich, alexander.raake}@tu-ilmenau.de

[9] SeeReal Technologies GmbH, Dresden, Germany, ez@seereal.com

[10] Schönhagen Airport, Trebbin, Germany, drschwahn@edaz.de

concepts such as edge computing introduced in 5G networks, enhance the clients with new capabilities making new types of server-client-splits possible. Sixth-generation networks are aiming to go even beyond this concept by flexibly distributing processing power over the whole backbone network. That is the core topic of the research project *6G NeXt – 6G Native Extensions for eXtended Reality (XR) Technologies*. Provided the tight integration with the communication infrastructure, the split computing approach would significantly enhance the 6G network experience. Based on two example use cases, *Smart Drones* anti-collision system and *Holographic Communication* (*HoloCom*), the project *6G NeXt* aims to develop and evaluate a joint connectivity and compute infrastructure, which will introduce new processing speeds in conjunction with highly dynamic geo-distributed computing capabilities within a mobile network.

Future outlooks for aviation anticipate an increasing number of remotely operated flight missions as well as an increasing number of manned flights [Rü22]. Until today, many strategies have been developed on the topic of how to implement unmanned aerial vehicles (UAVs) in the civil aviation airspace. In order to ensure a safe operation in shared airspace, a common anti-collision system will be necessary in the future. However, the key to success lies in the connectivity between different aircraft. In our work, we primarily focus on *Smart Drones*, a drone to drone anti-collision but future extensions to manned aviation are also conceivable.

The other application is supposed to bring the today's videoconferencing to the next level. By means of dedicated holographic displays, a real holographic 3D image can be created. Furthermore, the integrated eye-tracking capability allows the change of the view angle with no noticeable delay. This will provide people with a native 3D experience of their communication partners. However, the amount of holographic data is tremendous, which makes video processing on the client device impossible. Thus, the majority of processing steps such as rendering need to be performed on the side of the mobile network.

The remainder of the paper is as follows. In the next two sections, we provide a description of the targeted use cases. The appropriate architecture called to support our anti-collision system is shown in the Sect. 4. Here, we also provide a detailed description of technical solutions, required to support this architecture. Finally, Sect. 5 summarises the paper.

2 Smart Drones

The idea of an anti-collision system for drones or UAVs is derived from the research project VIGA (Virtual Instructor for General Aviation) [Rü22]. Here, the approach was based on a simulation of an aircraft that runs faster than the real life. In this way, evaluations of the possible flight trajectory can be estimated and warnings about hazardous situations can be provided. The proposed *Smart Drones* anti-collision system extends this approach to the UAVs. To do so, the application is split into two parts, which exchange data via a wireless link:

- Ground side: ground-based simulation engine with high processing power, which captures the UAVs' flight data and performs predictions of the flight paths.
- UAV side: flight controller, which keeps the communication to ground stations as well as controls the UAV accordingly

The flight path simulations are offloaded to the ground station in order to reduce the required computation power and thus the takeoff weight of the UAVs. The ground station is meant to collect all data streams of each individual UAV flying in the specific sector. The transmitted data contains intended mission profile data and UAV-specific data (flight model values) as well as altitude, position, course, and velocity information (state values), which are acquired on the drone at a high rate. The digital twin representation of the UAVs is fed by this UAV-specific information and the simulation is carried out with respect to the currently received state values. If a collision scenario is predicted by the system, evading trajectories will be calculated and simulated with the digital twins before the manoeuvre is transmitted and performed by the real UAV.

The performance of the system relies on the timeliness of the transmitted data. Thus, low latency and high reliability in the communication link are important since the overall time of data transmission, simulation, and maneuver execution has to be minimized. Moreover, the movement of the UAVs through different sectors requires dynamic data distribution. This requires a software architecture that is capable of allocating the UAV's data dynamically and handing over existing simulation data to other sectors.

To prove the feasibility of *Smart Drones* system, a real-life testbed setup is planned at Schönhagen Airfield (EDAZ). The airfield is located to the south of Berlin and Potsdam and it is one of the largest commercial airfields in Germany. Moreover, the airfield supports the research activities in the region. The long-term cooperation with universities and research institutes, such as German Aerospace Center (DLR), TU Berlin, and others, leads to newly emerging research and development projects. The Technical University of Applied Science of Wildau is a permanent cooperation partner, that maintains its own research aircraft. The most recent developmental step is the construction of an aviation safety centre with notable partners from the air traffic control and safety sectors. The centre should make a name for itself through its participation in national and international research activities, the provision of test environments, and simulations of threat scenarios for training and research, including the development and implementation of training modules.

3 Holographic Communication

Holographic communication is expected to become the next generation of video communication, with key players such as Ericsson and Google announcing their respective investments into the technology in recent years [Er23; La21]. While Ericsson aims for a device setup built around commodity hardware, Google's Starline "relies on custom-built hardware and highly specialized equipment". However, what both solutions have in common is the general processing pipeline for holographic communication, which can be applied to any concrete implementation. It consists of three steps: capturing, rendering, and display. In the capturing step, a 3D camera setup, which can consist of one or more cameras, is used to record the communication participant. In the rendering step, the resulting data is then processed into the 3D representation of the recorded participant, before being displayed on a device capable of playing back holographic video.

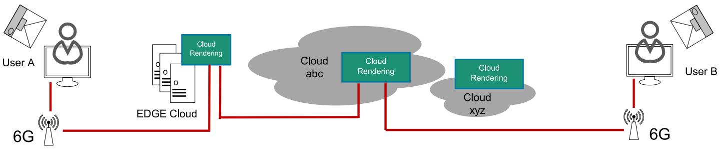

Fig. 1: *HoloCom* use case setup

The goal for *HoloCom* is the development of a real-time video communication system in a low-latency and high-bandwidth environment using specialized capturing and playback hardware to achieve an immersive and photo-realistic experience. Fig. 1 shows the exemplary setup of the use case. It details a split computing approach for the holographic communication pipeline, where capturing, rendering, and display are distributed between the communication participants and a local edge or cloud. When user A wants to call user B, image and depth information are continuously captured by a camera array on user A's side. The extremely high volume of data, potentially exceeding multiple terabytes per minute [Fr23], is transmitted via a 6G network to the nearest edge cloud, where it serves as input to a headless instance of the 3D engine Unity. In Unity, a 3D representation of user A is constructed from this data. Subsequently, the 3D representation is captured by a virtual camera, serving the data in a format required by the holographic display on user B's side. The display used mimics visual information used by the eyes to process depth in real-world objects, making the communication partner appear as they would in a face-to-face conversation. Enabled by this approach and coupled with eye tracking, the user is, within limits, free to move around in the physical space. The same pipeline is executed for user B, highlighting the necessity for high bandwidth and low latency, ensuring low delay without asynchronicity to achieve a good user experience.

4 System Architecture and Key Technologies

As described in previous sections, *Smart Drones* system as well as *HoloCom* are complex applications. On the one hand, they consist of mobile clients, which maintain wireless connection to the mobile network. On the other hand, the processing resources such as ground station or rendering facilities, are provided on the backbone network. Moreover, in order to cope with tight latency requirements, it is advantageous to split the processing tasks between edge clouds in the vicinity of the clients as well as high-performance centralised clouds.

Thus, resource orchestration requires tight coordination between communication and computing infrastructure, based on the behavior of the clients. In this paper, we provide a multi-layer architecture, which is called to enhance the inter-layer coordination. Whereas the architecture shown in Fig. 2 is generic and applies to both applications, its functionalities we will exemplary explain its functionalities based on the *Smart Drones* system.

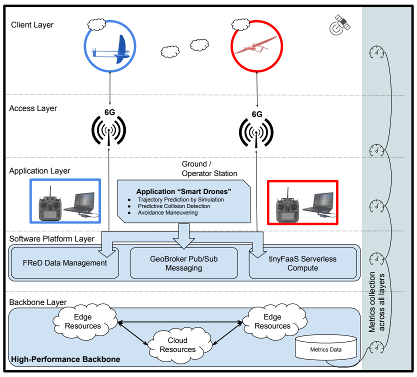

Fig. 2: Joint communication and compute architecture

This architecture will be able to provide the optimal resource allocation for all participants throughout the whole stack. Here, the UAVs are assigned to the *Client Layer*. They are logically connected to the ground control station, which is distributed on the *Application Layer* by means of mobile radio access network (RAN) at the *Access Layer*.

On the other hand, RAN maintains a physical connection to the *Backbone Layer*, where the high-performance backbone provides a convolution of communication network and computing infrastructure. In order to orchestrate the allocation of the applications to the processing resources, we introduce *Software Platform Layer*. Here, Function-as-a-Service (FaaS) based edge-cloud software platform provides an abstraction layer for the cloud-based applications running on *Application Layer*. The performance parameters of the functionalities on all layers are captured on the *Backbone Layer*. Here, optimisation of the resource allocation is performed. Taking into account the dynamically changing application environment, *Backbone Layer* anytime ensures appropriate resource allocation according to application demands.

4.1 Radio Access Network

In order to integrate the split computing infrastructure with the communication network, the latter should be able to flexibly react to the computation demands of the applications on the one side, and to the environmental influences on the quality of provided communication links on the other side. Therefore, it is necessary to address the inherent uncertainty and dynamic nature of the wireless environment and radio access network. While these requirements cannot always be guaranteed, machine learning (ML) models can be used to adapt to the spatiotemporal dynamics of wireless networks in advance. By providing predictive analytics to both end-user applications and radio access network elements, ML models can help to ensure the most efficient and effective operation of the system. Those are making it possible to make highly accurate predictions about changes in quality of service (QoS) and radio key performance indicators (KPIs), like radio environment maps (REMs), channel distribution information maps and spectral efficiency within the radio access network [Pe22; Zh22b]. By providing predicted QoS information, radio resource management schemes can offer more reliable future QoS guarantees to individual users, even when poor performance is expected. The characteristics of channel prediction methods on real communication data are also tested and verified by establishing an software-defined radio (SDR) based cellular communication and channel measurement [Zh22a].

Furthermore, the reliability of the communication channel can be increased by the adjustment of the coding schemes based on the predicted QoS values. In order to efficiently react to the changes of the channel conditions, the coding scheme should feature appropriate flexibility and agility. In our work, we investigate the utilisation of fountain codes to fulfil these requirements. A fountain encoder is able to generate an arbitrary number of codewords out of a payload packet, whereas only a certain number of them is required to decode the packet. Thus, the coding rate is not fixed, but can be adjusted “on the fly”. Since the fountain codes can withstand a drop of codeword, but a bit error, they should be utilised in conjunction with forward error correction (FEC). Even more, it was shown, that combination of error-detecting codes with fountain codes gains the performance of the communication system [Be08; Ka20]. The major advantage of fountain codes is their flexibility. Based on predicted channel quality, the coding rate can be flexibly adjusted in order to meet the QoS requirements [Me22]. Moreover, fountain coding schemes are capable of heterogeneous traffic requirements for different application types. Since a fountain encoder can be flexibly re-adjusted for any piece of payload, it will provide the agility to set up the optimal coding rate based on the traffic type as well as on the predicted channel conditions.

Final flight tests at Schönhagen Airport require an according wireless infrastructure. The modifications of wireless protocols will be realized and evaluated based on SDR platforms such as USRP. This concept is appropriate for ground-based infrastructure, whereas UAVs needs to be equipped with lightweight flight-capable terminal devices. For this purpose, the algorithms described above will be integrated with a dedicated communication module *NetMobilBox* [LW], the development of which goes back to the *5G NetMobil* project. Besides multi-connectivity capabilities, it provides the interfaces for the integration of additional sensors to capture positions and movement state information as well as communication to the UAV flight controller.

4.2 High-Performance 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 distribution of computing tasks, and the deep integration of artificial intelligence (AI) to optimize the overall system. The concept of a High-Performance Backbone is based on the ability to deploy the computation tasks at places, where sufficient network capabilities, CPU/GPU, and memory are available. In the following, the key technologies for a joint communication and computing infrastructure are presented.

4.2.1 Split Computing

Mobility-related services and applications can be found in the automotive-, rail- or drone industries. They are typical examples of distributed applications and split computing. The *Smart Drones* application illustrates how the proposed architecture supports a smart distribution of computing tasks. In this scenario, drone steering and control, trajectory prediction, predictive collision detection, and avoidance maneuvering tasks are distributed between drone, edge, and central cloud deployments as seen in Fig. 2.

Distribution of computation tasks in the backbone layer based on KPIs enhances the concept of edge computing and provides new possibilities for future applications. The core of this concept is to match the requirements of an application such as *Smart drones* anti-collision system with the most suitable resources and connections that are available based on KPIs. *3GPP* standardized in release 18 of its 5G specification an *Architecture for enabling Edge Applications*. 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.

4.2.2 Cross-Layer Metrics Function

Before a computing task can be assigned to a certain computing infrastructure, it needs distinct knowledge about available cloud resources and connectivity characteristics. A new proposed metric function aims at gathering metrics and measurement data from *Application*, *Software Platform*, and *Access Layer*, as shown in the Fig. 2 on the right side.

4.2.3 Discovery and Broker Service Function

A *Discovery Service* [Mic22] is a typical part of a microservice architecture. Clients running in different locations or changing their locations need to find their optimal endpoint to connect to. The *Discovery Service* provides e. g. the UAVs clients with the network address of an anti-collision system service. Additionally, a service broker among others helps cloud-based services to find free cloud capacities as well as also to start a new instance based on requirements described with KPIs. When a service recognizes a decreasing service quality due to missing memory or CPU/GPU in the current cloud environment, it might ask the *Discovery* as well as *Broker Service 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. Second, it helps the service to deploy and start a new service instance using e. g. Docker and automated infrastructure tools.

In order to achieve the best suitable network connectivity, it is important to understand the nature of the traffic, e. g. in the surrounding of an aerodrome. The most crucial traffic parts must be identified and treated with a higher priority. This may be achieved by routing it to dedicated network slices with inherently higher traffic prioritisation, potentially combined with dedicated and reserved radio resources for such slices. As an additional mechanism, individual traffic flows can be elevated to a higher priority level on the fly by using QoS related application programming interfaces (APIs) as specified in the CAMARA Telco Global API Alliance.

4.3 Edge-Cloud Software Platform

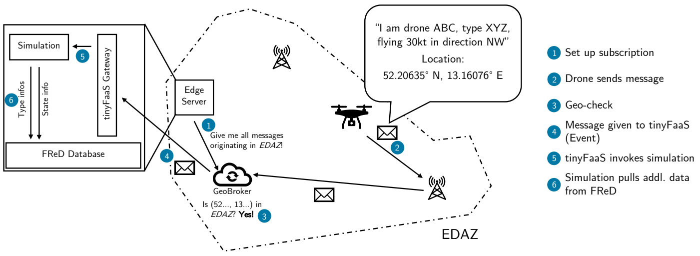

Fig. 3: A workflow of the drone anti-collision system using serverless abstractions for the edge-cloud.

In order to make the high-performance communication and computation backbone suitable for applications, we need to use novel software abstractions. A serverless abstraction level can provide a unified interface to geo-distributed cloud and edge resources managing computation, messaging, and data replication.

We rely on distributed Publish/Subscribe (Pub/Sub) to provide messaging across distributed services, clients, and applications in the edge-to-cloud serverless platform. We use the distributed *GeoBroker* [HB20] that extends Pub/Sub with geographical context: *GeoBroker* can automatically filter messages not only by topic but also by geographic relevance so that subscribers receive only messages published in their proximity. This introduces a natural geographic sharing of responsibilities among replicas of an application service.

Serverless computation is implemented using the *tinyFaaS* [PB20]. It is a lightweight implementation of the FaaS paradigm, where applications are composed of interconnected stateless functions that are invoked in an event-driven manner.

Data replication is a key concern in distributed edge-to-cloud platforms: Always using a central cloud replica of a database incurs a high access latency for edge services. *FReD* is a data management middleware for the edge-to-cloud continuum that manages data replication for applications [HGB20]. Using geographically distributed nodes, applications simply specify a set of locations to replica a table of data to.

In Fig. 3, it is shown, how these abstractions can support the *Smart Drones* anti-collision system. The edge service simply specifies a geographic context that it deems relevant, e. g., the *EDAZ* airport perimeter. Any drone can send its location and status updates to a *GeoBroker* instance without specific knowledge on *which* service is interested in this data and *where* an instance of this service is running. *GeoBroker* then forwards this message to the edge service after confirming its topical and geographical relevance (with a check, whether the origin location lies in the specified geo-fence). The edge service lets *tinyFaaS* invoke the simulation component in an event-driven manner: *if* a message is received from a drone, *then* execute the simulation. A simulation instance is then automatically started by *tinyFaaS*. This instance can access a local copy of relevant data through an interface to *FReD*, e. g., to retrieve information about the maneuvering capability of a type of drone. It completes its calculations and returns instructions to the drone if necessary.

5 Conclusion

The 5th generation of mobile networks is still in the roll-out phase. However, the technologies introduced there, such as the edge computing concept, show certain limitations. Their capability will be not sufficient to serve a number of future-oriented applications, which are currently emerging. The UAVs control and anti-collision system *Smart Drones* and the *HoloCom* communication system, described in this paper, are just some of them. On the one hand, these applications set high demands on the capability of the wireless link to the clients, such as high reliability and low latency. On the other hand, a sophisticated distributed processing engine on the backbone network is required, which should be flexible to follow the mobile clients on their track. In the future 6G networks, communication and processing cannot be considered separately, but they will grow together into a joint communication-and-computing infrastructure.

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

[Be08] Berger, C. R. et al.: Optimizing Joint Erasure- and Error-Correction Coding for Wireless Packet Transmissions. IEEE Transactions on Wireless Communications 7/11, pp. 4586–4595, Nov. 2008.

[Er23] Ericsson Holographic Communication, 2023, URL: `https://www.ericsson.com/en/ericsson-one/holographic-communication`, visited on: 10/06/2023.

[Fr23] Friese, I. et al.: True 3D Holography: A Communication Service of Tomorrow and Its Requirements for a New Converged Cloud and Network Architecture on the Path to 6G. In: International Conference on 6G Networking (6GNet). IEEE, 2023.

[HB20] Hasenburg, J.; Bermbach, D.: GeoBroker: Leveraging Geo-Context for IoT Data Distribution. Elsevier Computer Communications 151/, pp. 473–484, Feb. 2020.

[HGB20] Hasenburg, J.; Grambow, M.; Bermbach, D.: Towards A Replication Service for Data-Intensive Fog Applications. In: Proceedings of the 35th ACM Symposium on Applied Computing, Posters Track (SAC '20). ACM, Brno, Czech Republic, pp. 267–270, Mar. 2020.

[Ka20] Karrenbauer, M. et al.: A Study on the Application of Rateless Coding in Non-Cellular MIMO Systems for Machine-Type Communication. IFAC 53/2, pp. 8243–8248, 2020.

[La21] Lawrence, J. et al.: Project Starline: A high-fidelity telepresence system. ACM Transactions on Graphics (Proc. SIGGRAPH Asia) 40(6)/, 2021.

[LW] Logic Way Kommunikationsmodule, URL: `https://logicway.de/pages/dimmcpu.shtml`, visited on: 03/28/2023.

[Me22] Melnyk, S. et al.: Wireless Industrial Communication and Control System: AI Assisted Blind Spot Detection-and-Avoidance for AGVs. In: IN4PL. Pp. 1301–1306, Mar. 2022.

[Mic22] Service Discovery in Microservices, Nov. 2022, URL: `https://www.baeldung.com/cs/service-discovery-microservices/`, visited on: 04/26/2023.

[PB20] Pfandzelter, T.; Bermbach, D.: tinyFaaS: A Lightweight FaaS Platform for Edge Environments. In: Proceedings of the Second IEEE International Conference on Fog Computing (ICFC 2020). IEEE, Sydney, NSW, Australia, pp. 17–24, Apr. 2020.

[Pe22] Perdomo, J. et al.: QoS Prediction-based Radio Resource Management. In: 2022 IEEE 96th Vehicular Technology Conference (VTC2022-Fall). Pp. 1–6, 2022.

[Rü22] Rüther-Kindel, W. et al.: VIGA — Virtual Instructor for General Aviation. In: CEAS EuroGNC 2022. May 2022.

[Zh22a] Zhou, Q. et al.: Deep Learning-Based Signal-to-Noise Ratio Prediction for Realistic Wireless Communication. In: 2022 IEEE 95th Vehicular Technology Conference: (VTC2022-Spring). Pp. 1–5, 2022.

[Zh22b] Zhou, Q. et al.: Performance Evaluation over DL-Based Channel Prediction Algorithm on Realistic CSI. In: 2022 IEEE 96th Vehicular Technology Conference (VTC2022-Fall). Pp. 1–5, 2022.