Resources Allocation for Hybrid Cloud-Edge Computing in 5G Network Slicing

Mlungisi Molefe[1], Khulekani Sibiya[2], Bakhe Nleya[3]

Department of Electronic and Computer Engineering, Durban University of Technology
[1]mlungisim3@dut.ac.za,[2]21345559@dut4life.ac.za,[3]bakhen@dut.ac.za
**Durban, South Africa*

***Abstract*— In typical heterogeneous networks such as 5G and beyond, innovative technologies such as edge computing and network slicing can enhance overall network performance in terms of handling critical mission services as they often require extremely low latencies. Notably, network slicing facilitates the provisioning of virtual slices with different characteristics to serve different end-user requirements. The Network operator achieves this goal by utilizing the already existing physical wireless network resource. Current resource provisioning schemes suffer inadequacies in scalability and flexibility Thus to support both Cloud and Edge Computing in 5G and beyond networking, the work herein proposes a novel low latency scheme that affords dynamic and intelligent allocation of multi-dimensional resources. It bases on a Hybrid Cloud-edge Network Slicing (HCENS) architecture on leveraging both Cloud and Edge Computing The proposed scheme creates a flexible, scalable as well as energy efficient resource provisioning. Its architecture comprises both centralized units (CUs) and distributed units (DUs). These provide storage, that in turn enhances function partitioning for various network slices. Several agent-based simulations scenarios are caried out in evaluating the efficacy of the proposed scheme. Obtained analytical and simulation results indicate drastic reductions in network latencies for critical mission end user services. This couples with reductions in storage requirements.**

***Keywords*— Network slicing, Edge computing, Cloud, Latency, 5G, Resources Allocation.**

I. INTRODUCTION

With emerging technologies such as internet of things (IoT), smart homes, virtual reality and autonomous vehicles, network demands continuously evolve towards deeper heterogeneity and diversification. The planned gradual deployment of 5G as a stand-alone network will in the interim allow its coexistence with legacy generations such as 4G LTE. In 5G non-standalone mode, 4G LTE base stations can be viewed as signalling anchors, while the recently deployed 5G base stations can be perceived or rather viewed as user path data points. Gradually, a fully fletched 5G standalone network deployment will be completed in which it will act as the base cloud-based core. This accomplishment will imply 5G base stations becoming signalling anchors. The International Telecommunication Unit (ITU) has categorised 5G service use cases into three categories, namely, ultra-reliable and Low Latency Communication (uRLLC), enhanced Mobile Broadband (eMBB) and limitless connectivity or massive Machine Type Communication (mMTC). In the uRLLC category, services are arranged to satisfy low latency, fail safe, and security sensitive applications and services. Critical mission applications and services, e.g., telemedicine, and self-driving automobiles will fall in this category. The eMBB category, services are arranged for enhanced connection at cell edges, thus, to accommodate large bandwidth services with typical downlink and uplink transmission speeds in the order of 1Gbps or more. Example services and applications in this category include high-definition video streaming and virtual reality. The mMTC category, generally suites low-power IoT devices. Inadequacies with capabilities of current and legacy networks, means that in meeting the demands of the 5G network and beyond, network slicing based architectures coupled with Edge Computing techniques is a viable design solution. Perhaps, to achieve enhanced scalability in resource provisioning, whisking away current network slicing techniques towards the edge in the envisaged 5G and beyond networks would be quite helpful.. In that way resource coordination and provisioning would be carried out at both the core and edge. As already established Edge computing efficiently offers support for low latency for applications, services and for user requests whereas, cloud level computing provides additional resource for edge devices that may require more than the limited already at their disposal. In realization, the edge-cloud integrated systems comprise extensive heterogeneous hardware systems and base stations. It generally problematic to harmoniously integrate these systems given their being managed within frameworks of differing interfaces and protocols in present 5G network slicing scenarios. Both network functions virtualization (NFV) and software-defined networking (SDN) can aid in strategically integrating edge computing and cloud level sub-slices in future 5G and beyond networks. In short to guarantee flexibility, programmability as well as reconfigurability both virtualization and softwarization will be implemented in order to enhance the sharing of resources by billions of devices simultaneously in the small physical space.

The paper proposes a hybrid cloud-edge computing architecture for 5G networks and beyond that is flexible, agile, scalable, and reliable to cater for big data and low latency transmission for different network slices. The rest of the paper is organized as follows: Section II discusses the related literature on the topic, Section III proposes a hybrid cloud-edge network slicing (HCENS) architecture for 5G and beyond networks, section IV discusses the resources allocation model for HCENS architecture and suggesting appropriate network slicing and resources allocation algorithm, section V evaluates the performance of the proposed HCENS architecture, and finally Section VI concludes the paper.

II. RELATED WORKS

Research in 5G and beyond networks resource allocation and optimization has been conducted extensively. The focus is on network-slicing techniques as they are a promising solution to the o enhancement of both sound resource management as well as utilization. The work in [1] extensively explores this same area, in particular possible challenges as well as future trends. The work in [2] proposes and evaluates a novel heterogeneous algorithm that seeks to enhance both QoE and fairness to all end-users, whilst at the same time minimizing risks of interference among traffic beams. System reliability, resilience, robustness, and interference efficiency are explored in [3], in which case a heterogeneous environment is assumed. Resource allocation optimization and network resilience energy efficiency problems are explored in [4]. The same work also takes both security and privacy into consideration. Spectrum scarcity in current networks prompts the authors in [5], [6] to further explore resource allocation in the 5G wireless environment. Mobile edge computing is investigated in various works as a candidate solution for addressing the latency problems experienced by various applications and services that are of critical nature in heterogeneous 5G network environments. Similarly works have expanded towards addressing efficiency in radio resource management in virtual networks. Recently, most literature tend to broadly categorize network slicing into infrastructure versus Resource spectrum-based slicing. E.g., the authors in [7] propose a profit-aware joint power allocation and slice resource allocation considering the core network's capacity as well as end users' QoE expectations. In this same work, the emphasis is on enhancing network throughput in multi-slices and multi-user cases. The scheme proposed herein, though demonstrating improved performances, however, tends to burden the central controller as each individual slice must allocate resources individually. In [8], the authors explore a method of enhancing hardware resource distribution for edge nodes in a multitier MEC hierarchy. In addition to a centralized unit, they also consider dispersed units with edge nodes that have varying processing capacities. To boost the total computing capability of a 5G-based MEC system, a parametric Bayesian optimizer is designed for hardware resource allocation. The scheme is evaluated and found to outperform pseudorandom resource allocation in terms of the percentage of performed computing jobs under the specified budget limitations.

In [9], the authors indicate various RAN slicing algorithms and assess them in terms of the level of detail employed in radio resource assignment as well as the obtained isolation and customization to make a proposal for a network virtualization substrate that is built on a slice scheduler that distributes radio resources to the tenants and a flow scheduler that plans data transmission for the users in each network slice. The authors further propose resource allocation in RAN by placing multiple DU and CU in the 5G network slicing, this allows the network to process the data in the three slices in an effective and efficient manner.

The work in [10] suggests a unified radio access network (RAN) slice service provisioning architecture that will enhance overall network performance by allocating the bare minimum of requested bandwidth. In the same work, the effect of bandwidth allocation on network service quality is solely considered when it comes to bandwidth distribution between slices. From an end-to-end viewpoint, they do not execute differentiated slice bandwidth allocation, and neither are network operational costs considered when provisioning bandwidth. For storage and better function partitioning for different network slices, our HCENS architecture is designed with deployments of centralized units (CUs) and distributed units (DUs). Our HCENS design gives the lowest latency and completes all successful runs utilizing a variety of processing and storage resources because it exceeds their slice service provisioning architecture network models in terms of latency and storage capacity.

Recently standardization initiatives in regard to network slicing have been carried out by various groups including; the SFII (Slicing Future Internet Infrastructures) [15], NECOS (Novel Enablers for Cloud Slicing) [16], SELFNET [17], and MATILDA [18], among others. These initiatives mostly address technological elements, architectures, and slicing techniques.

This slicing design, initiatives, however, merely address the conceptual and analytical simulation of the fundamental structures and functions that make up the slicing process in a preliminary manner or did not even mention them as issues to be resolved in the future. The conceptual and analytical modeling of communication slices is a fresh addition to the field of network slicing. Thus, our HCENS design provides the lowest latency and completes all successful runs taking advantage of a range of processing and storage resources because it exceeds their slice service provisioning architecture network models in terms of latency and storage capabilities.

III. HYBRID CLOUD-EDGE NETWORK SLICING ARCHITECTURE

In this section, we propose a hybrid network slicing and cloud-edge computing which we term Hybrid Cloud-Edge Network Slicing (HCENS). It is hybrid in the sense that it executes computing functions in more than a single computing platform. This implies that the architecture overall comprises a dual layer set of terminal layer configurations, one distributed and the other centralized. The earlier is composed of 5G distributed units (5G-DU) which together serve as an access point for end users. This layer is key in rendering edge computing functionalities; hence we refer to it herein as Distributed Multi-access Edge Computing (D-MEC). The centralized layer comprises several 5G centralized units (5G-CU) and is termed the Centralized Multi-access Edge Computing (C-MEC) layer. It does render limited edge computing as well. The HCENS architecture is illustrated in Fig. 1.

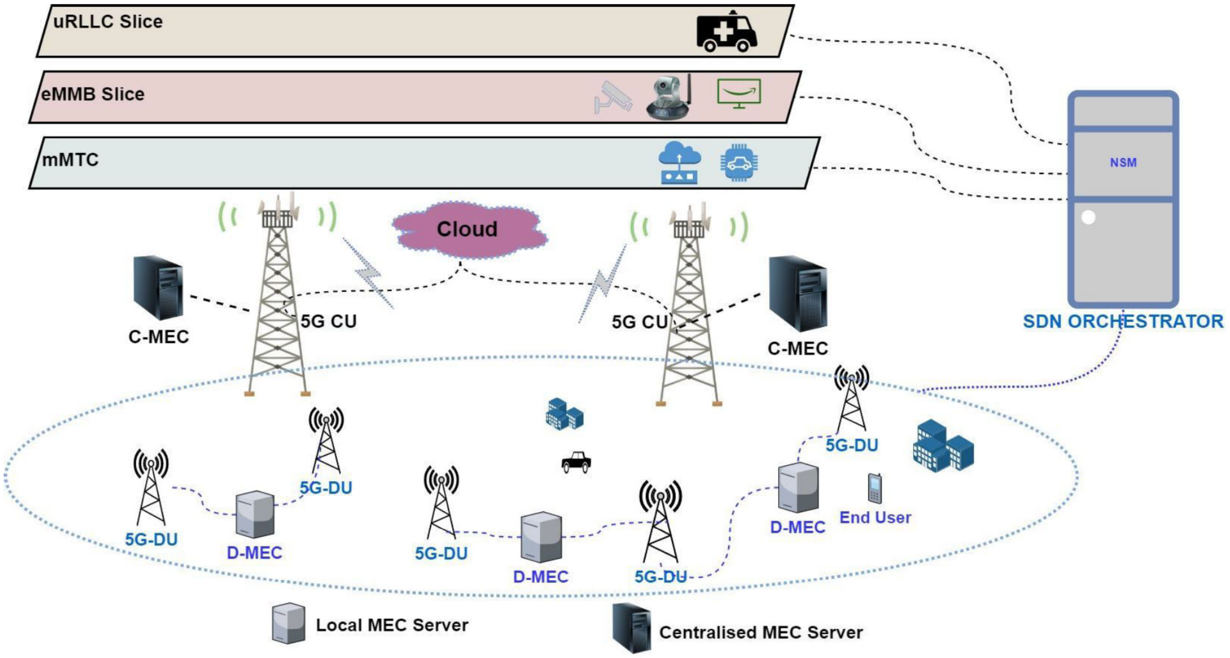

Figure 1. HCENS architecture

Its overall architecture comprises six agents; cloud, 5G-CU, C-MEC, 5G-DU, D-MEC, and end users. The deployment of distributed units are significant to provide more convenient function partitioning for different network slices and more scalable access network updating. The 5G-CU and 5G-DU base stations can be viewed as logical nodes in which 5G-DU nodes are operating in the lower layer or closer to end users, while 5G-CU nodes on the other hand play a significant role in controlling functions, such as session management (SMF), mobility (MF), data transmission (DTF) and radio access network. The D-MEC is a type of MEC that is like fog computing in terms of computing or processing power, and it is designed with tools for delivering services to end users. Likewise, C-MEC is another type of MEC that has higher computing power and high storage capabilities in comparison to D-MEC.

The HCENS architecture further leverages network slicing and software-defined networking (SDN) to allocate network resources effectively and improve the quality of service (QoS). The resource allocation amongst the three network slices is achieved by the adoption of SDN for centralized management and control. The SDN orchestrator is requested to allocate a slice to a specific service, and it performs resource allocation calculations to select the appropriate slice. Moreover, the HCENS architecture adopts optical front haul and optical backhaul using passive optical networking. Since network slices are logical networks, therefore this means that some previous hardware functions have been abstracted through the adoption of network function virtualization (NFV). This architecture stands out compared to other architectures in literature because it offers a two-layer edge computing depending on the amount of computed or processed data, it also offers network slices that enforce end-to-end service level agreements by delivering and maintaining a specific QoS across the radio access network, transport network, and the backbone core network.

A. Slice Creation Across the Network

The SDN orchestrator has an internal Network Slice Manager (NSM) that acts as an end-to-end network slice orchestrator. The NSM's role is to dynamically categorize sets of network functions that satisfy some pre-defined service characteristics. These service characteristics can contain a mixture of features, such as latency, transmission speed, location, and priority.

The network slice creation follows three basic steps as follows:

i. Identifying the needed network functions, assigning the necessary resources, as well as generating the network functions.
ii. Determining the actual location o the network functions (core/ edge) and interconnecting them.
iii. Registering details of the slice as well as mapping the requested resources to the user.

Furthermore, the NSM rearranges the network resources by supporting auto-scaling. In this case, the network slice resource is resized to match the actual demand.

B. Slice Allocation to End-Users

The assignment of a slice to the requesting end user is as follows:

- End user initializes (or initiates) a service that requires a network slice from the NSM.
- A network slice selection assistance information (NSSAI) that relates to the service needs is acknowledged to the end user. Finally, the 5G network generates a network slice based on the information received on the

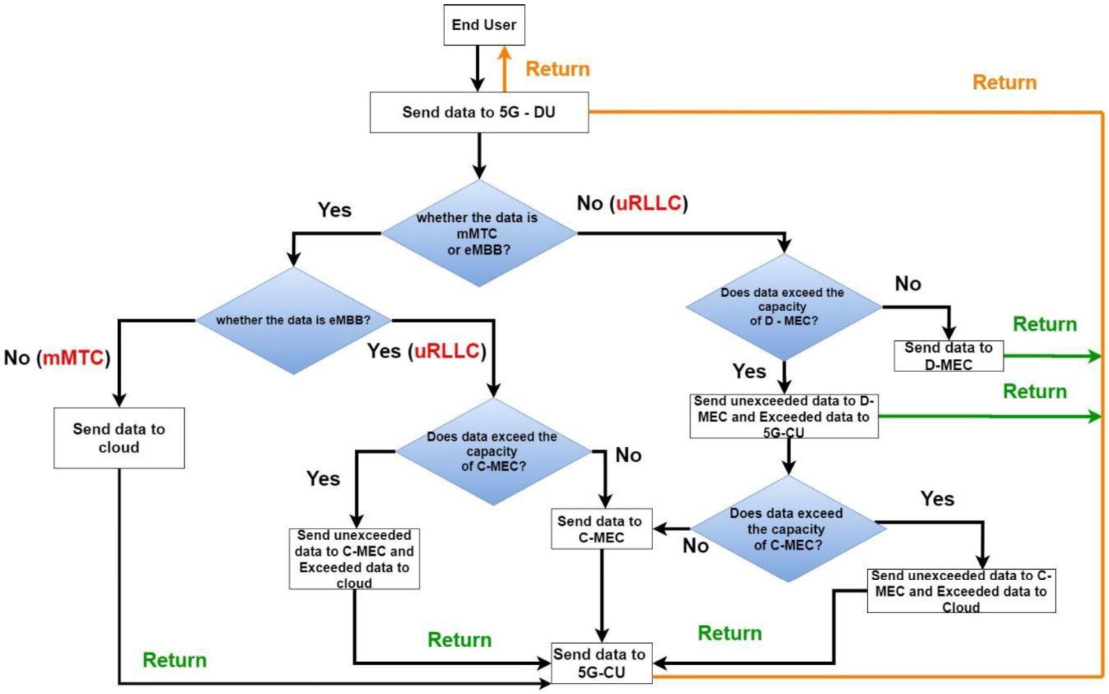

Figure 2: Resources allocation and network slicing flowchart

NSSAI for serving the end user. Assuming an end user requires an uRLLC network slice, the NSM recognizes the need to allocate a slice to the end user, and dedicated resources are assigned to the slice, such as ultra-low latency and high-capacity resource pooling.

IV. RESOURCES ALLOCATION

In this section, we briefly outline the resource allocation and network slicing algorithms as executed by the SDM orchestrator. It requires primitives such as latency and capacity demands. The algorithms combined are illustrated in Fig.2.

A. *Latency Evaluation*

We have chosen the maximum data storage of D-MEC and C-MEC be 1GB and 2GB respectively. The total latency (L_T) for a group of data packets (PD_i) ($i \epsilon$ 1,2,…N) is equivalent to the sum of maximum latency between the cloud (L_{cloud}), C-MEC ($L_{C\text{-}MEC}$), and D-MEC ($L_{D\text{-}MEC}$), as expressed by the following :

$$L_T = \sum_{i=1}^{N} max\ PD_i\{L_{cloud}, L_{C-MEC}, L_{D-MEC}\} \tag{1}$$

where,

$$L_{D-MEC} = 2\left(\frac{V_{Data}}{\alpha*S_{EU}} + \frac{V_{D-MEC}}{S_{5G-DU}}\right) + \frac{V_{D-MEC}}{C_{D-MEC}} \tag{2a}$$

$$L_{C-MEC} = 2\left(\frac{V_{Data}}{\alpha*S_{EU}} + \frac{V_{C-MEC}}{S_{5G-CU}}\right) + \frac{V_{C-MEC}}{C_{C-MEC}} \tag{2b}$$

$$L_{cloud} = 2\left(\frac{V_{Data}}{\alpha*S_{EU}} + ln\ ln\ H * \frac{V_{cloud}}{S_{5G-CU}}\right) + \frac{V_{cloud}}{C_{cloud}} \tag{2c}$$

where, V_{Data} , V_{D-MEC}, V_{C-MEC}, and V_{cloud} are the volumes of data from the end user (EU) to the 5G-DU, from 5G-DU to D-MEC, from 5G-CU to C-MEC, and from 5G-CU to cloud, respectively. L_{D-MEC}, L_{C-MEC}, and L_{cloud} are the total latency for D-MEC, C-MEC, and cloud, respectively. S_{EU}, S_{5G-CU}, S_{5G-DU} are the speed rates of EU, 5G-CU, and 5G-DU, respectively. C_{cloud}, C_{C-MEC}, and C_{D-MEC} are the computational power of cloud, C-MEC, and D-MEC, respectively. The parameters H=10 and α= [0.9, 1] are kept constant,.
Consequently the aggregate data volume is expressed as;.

$$V_{Data} = V_{C-MEC} + V_{D-MEC} + V_{cloud} \tag{3}$$

B. *Resources allocation for uRLLC network slice*

For resource allocation in the uRLLC , the following set of equations will apply;

$$\phi = \frac{V_{Data}}{a*S_{EU}} + \frac{V_{D-MEC}}{S_{5G-DU}} + \frac{V_{Data}-V_{D-MEC}}{S_{5G-CU}} \tag{4}$$

$$\eta = \frac{V_{Data}}{a*S_{EU}} + \frac{V_{D-MEC}}{S_{5G-DU}} + \frac{V_{C-MEC}}{S_{5G-CU}} + \ln H * \frac{V_{Data} - V_{D-MEC} - V_{C-MEC}}{S_{5G-CU}} \qquad 5)$$

$$\mu = \frac{V_{Data} - V_{D-MEC} - V_{C-MEC}}{S_{5G-CU}} \tag{6}$$

$$\lambda = \frac{V_{Data}-V_{D-MEC}}{C_{D-MEC}} \tag{7}$$

Consequently, the end-to-end latency in the uRLLC slice (LT(u)) is computed from;

$$L_{T(u)} = \begin{cases} 2\left(\frac{V_{Data}}{a * S_{EU}} + \frac{V_{Data}}{S_{5G-DU}}\right) + \frac{V_{Data}}{C_{D-MEC}}; & V_{Data} \leq V_{D-MEC} \\ 2\,\phi + \frac{V_{D-MEC}}{C_{D-MEC}} + \lambda; & V_{D-MEC} < V_{Data} \leq V_{D-MEC} + V_{C-MEC} \\ 2\,\eta + \frac{V_{D-MEC}}{C_{D-MEC}} + \frac{V_{C-MEC}}{C_{C-MEC}} + \mu; & V_{Data} \geq V_{D-MEC} + V_{C-MEC} \end{cases} \quad (8)$$

C. Resources allocation for eMBB network slice

For resource allocation in the following equations will apply;

$$\beta = \frac{V_{Data}}{a * S_{EU}} + \frac{V_{Data}}{S_{5G-CU}} \quad (9)$$

$$\delta = \frac{V_{Data}}{a * S_{EU}} + \frac{V_{C-MEC}}{S_{5G-CU}} \quad (10)$$

$$\gamma = \frac{V_{Data} - V_{C-MEC}}{S_{5G-CU}} \quad (11)$$

$$\theta = \frac{V_{Data} - V_{C-MEC}}{C_{cloud}} \quad (12)$$

Finally, the end-to-end latency in the eMBB slice ($L_{T(e)}$) is given by;

$$L_{T(e)} = \begin{cases} 2\beta + \frac{V_{Data}}{C_{C-MEC}}; & V_{Data} \leq V_{C-MEC} \\ 2(\delta + \ln H * \gamma) + \frac{V_{C-MEC}}{C_{C-MEC}} + \theta; & V_{Data} \geq V_{C-MEC} \end{cases} \quad (13$$

D. Resources allocation for mMTC network slice

Latencies are calculated according to the following set of equations;

$$L_{T(m)} = 2\left(\frac{V_{Data}}{a * S_{EU}} + \frac{V_{Data}}{S_{5G-CU}} + ln\, ln\, H * \frac{V_{Data}}{S_{5G-CU}}\right) + \frac{V_{Data}}{C_{Cloud}} \quad (14)$$

$$L_{T(u)} - L_{T(e)} = 2\left(\frac{V_{Data}}{S_{5G-DU}} - \frac{V_{Data}}{S_{5G-CU}}\right) + \frac{V_{Data}}{C_{D-MEC}} - \frac{V_{Data}}{C_{C-MEC}} \leq 0 \quad (15)$$

$$L_{T(e)} - L_{T(m)} = -2\, ln\, ln\, H * \frac{V_{Data}}{S_{5G-CU}} + \left(\frac{V_{Data}}{C_{D-MEC}} - \frac{V_{Data}}{C_{Cloud}}\right) \leq 0 \quad (16)$$

V. PERFORMANCE EVALUATION

In this section we evaluate the overall performance of the HCENS versus the traditional MEC. Fig. 3, plots our simulated results in which it is clear, our architecture model outperforms. The end-to-end latency evaluation takes into account storage capacities at different edges.

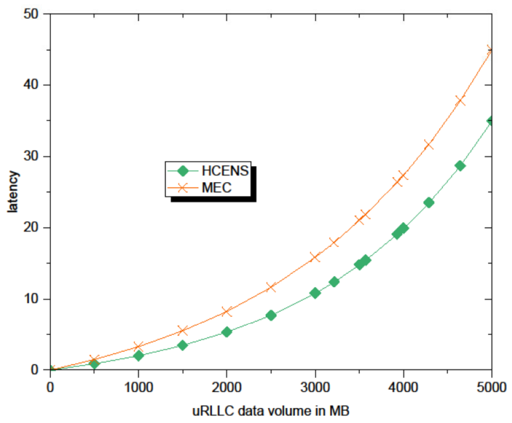

Figure 3: Latency for HCENS vs MEC in the uRLLC Slice

Results for the latency evaluation for the eMBB network slicing case is represented in Fig. 4. Once again by comparisons, the HCENS architecture outperforms.

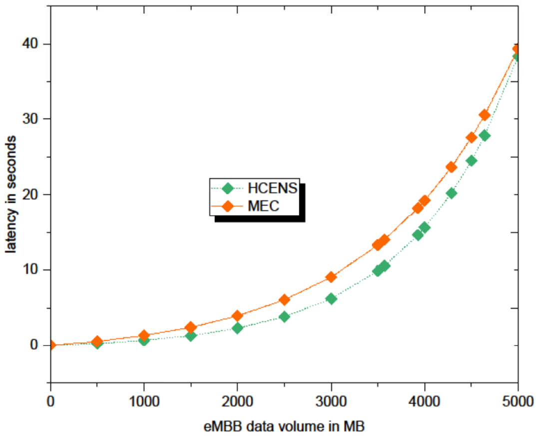

Figure 4: Latency for HCENS vs MEC in the eMBB Slice

We further explore the latency at different storage capabilities on the D-MEC of the HCENS model. In this case work, we assume that the cloud servers have 2.5 times more storage capability compared to D-MEC servers, on the other hand, C-MEC servers have about 1.5 times more storage capability compared to D-MEC servers, hence the power consumption at these edges are also different depending on the amount of storage capability. The simulations on Fig. 5 and Fig. 6 show that the latency of different slices will decrease as the storage capacity increase. In Fig. 5, the latency of eMBB and uRLLC are less than the latency of mMTC because eMBB and uRLLC uses edge storages and processing while mMTC uses cloud storages and processing. Furthermore, in Fig. 6, the latency of the eMBB and uRLLC slices have the lowest latency at higher storage capability, while the mMTC slice remains relatively fixed because it only uses cloud storage and does not use any edge storage.

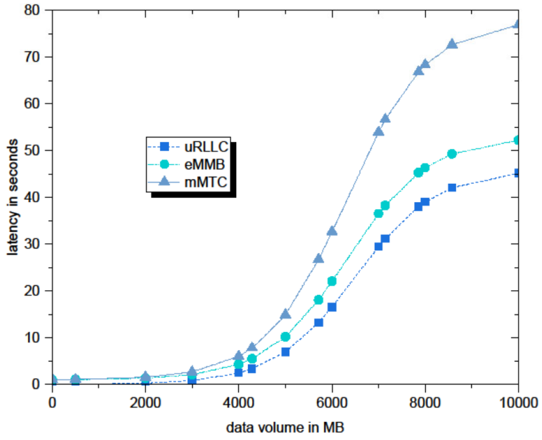

Figure 5: Storage capability of D-MEC = 1000 MB

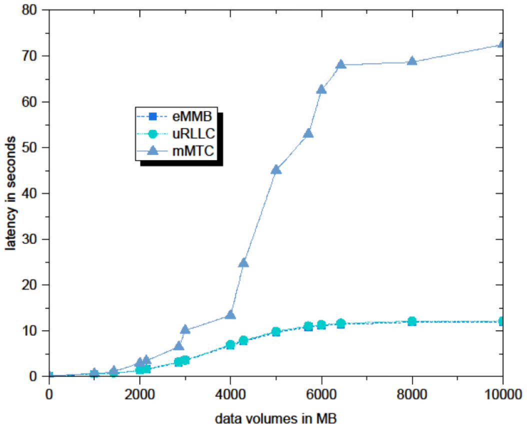

Figure 6: Storage capability of D-MEC = 6000 MB

VI. CONCLUSION

The paper proposed and evaluated low latency scheme that affords dynamic and intelligent allocation of multi-dimensional resources. It bases on a Hybrid Cloud-edge Network Slicing (HCENS) architecture on leveraging both Cloud and Edge Computing The proposed scheme creates a flexible, scalable as well as energy efficient resource provisioning. Its architecture comprises both centralized units (CUs) and distributed units (DUs). These provide storage, that in turn enhances function partitioning for various network slices. Several agent-based simulations scenarios are caried out in evaluating the efficacy of the proposed scheme. Obtained analytical and simulation results indicate drastic reductions in network latencies for critical mission end user services. This couples with reductions in storage requirements.

In this paper we presented a novel combination of edge computing, cloud computing and network slicing architecture which leverages on distributed units (DUs) for edge processing and storage, also leveraging on centralised units (CUs) for centralized storage and processing and cloud storage/processing for high volume data, we proposed an innovative system architecture referred to as the Hybrid Cloud-Edge Network Slicing (HCENS) architecture. Furthermore, we proposed a resources allocation model for different network slices (uRLLC, eMBB, and mMTC), to improve different user-oriented quality of services (QoSs). We conducted several tests to evaluate the latency performance of the HCENS model compared with the general multi-access edge (MEC) computing for different network slices. Finally, we evaluated the latency of each network slice under different storage capability in the distributed servers. The results suggest that our proposed HCENS model outperforms the MEC deployment in terms of latency for different network slices, and the advantage of using edge computing is significant on reducing latency as observed in the simulated results.

REFERENCES

[1] Y. Xu, H. Xie, and R. Q. Hu, ``Max-min beamforming design for heterogeneous networks with hardware impairments," IEEE Communication. Letters., Vol. 25, no. 4, pp. 1328_1332, Apr. 2021.

[2] Y. Xu, H. Xie, and R. Q. Hu, ``Max-min beamforming design for heterogeneous networks with hardware impairments," IEEE Communication. Letters., Vol. 25, no. 4, pp. 1328_1332, Apr. 2021.

[3] Y. Xu, G. Gui, T. Ohtsuki, H. Gacanin, B. Adebisi, H. Sari, and F. Adachi,``Robust resource allocation for two-tier HetNets: An interference efficiency perspective," IEEE Trans. Green Communication. Networks., vol. 5, no. 3,pp. 1514_1528, Sep. 2021.

[4] Y. Xu, H. Xie, C. Liang, and F. R. Yu, ``Robust secure energy efficiency optimization in SWIPT-aided heterogeneous networks with a nonlinear energy harvesting model," IEEE Internet Things J., early access, Apr. 13, 2021, doi: 10.1109/JIOT.2021.3072965.

[5] H.Wang, S.-H. Leung, and R. Song, ``Uplink area spectral efficiency analysis for multichannel heterogeneous cellular networks with interference coordination," IEEE Access, vol. 6, pp. 14485_14497, 2018.

[6] Z. Tan, X. Li, F. R. Yu, L. Chen, H. Ji, and V. C. M. Leung ``Joint access selection and resource allocation in cache-enabled HCNs with D2Dcommunications," in Proc. IEEE Wireless Communication. Networks. Conf. (WCNC),Mar. 2017, pp. 1_6.

[7] T. LeAnh, N. H. Tran, D. T. Ngo, and C. S. Hong, ``Resource allocation for virtualized wireless networks with backhaul constraints," IEEE Communication Letters., vol. 21, no. 1, pp. 148151, Jan. 2016.

[8] S. Zhang, "An overview of network slicing for 5G", IEEE *wirel Commun*. 2019;26(3):111-117.

[9] W. Saleh and S. Chowdhury, "RANSlicing: Towards Multi-Tenancy In 5G Radio Access Networks", International Journal of Wireless & Mobile Networks (IJWMN), Vol.14, No.2, April 2022 DOI:10.5121/ijwmn.2022.14204 43

[10] T. X. Tran, A. Hajisami, P. Pandey, and D. Pompili, "Collaborative Mobile Edge Computing in 5G Networks: New Paradigms, Scenarios, and Challenges", IEEE Communications Magazine, April 2017. Pp 54 -61.

[11] L. Feng, W. LI, Y. Lin, L. Zhu, S. Guo, and Z. Zhen, "Joint Computation Offloading and URLLC Resource Allocation for Collaborative MEC Assisted Cellular-V2X Networks", Special section on communication and fog/edge computing towards intelligence connected vehicles (ICVS), Digital Object Identifier 10.1109/ACCESS.2020.2970750.

[12] E. Šlapak, J. Gazda, W Guo, T. Maksymyuk, and M. Dohler, "Cost-Effective Resource Allocation for Multitier Mobile Edge Computing in 5G Mobile Networks", Digital Object Identifier 10.1109/ACCESS.2021.3059029.

[13] S. Huang, B. Guo, Y. Liu, "5G-oriented optical underlay network slicing technology and challenges", IEEE *Commun Mag*. 2020.

[14] J. Zhang, Y. Xiao, D. Song, L. Bai, Y. Ji,."Joint wavelength, antenna, and radio resource block allocation for massive MIMO enabled beamforming in a TWDM-PON based fronthaul", J Light Technol. 2019;37(4):1396-1407.

[15] S. B. Martins, C. T. Carvalho, F. Silva, and R. Moreira. "SFI2 Network Slicing Reference Architecture". Technical Report TR03/2022. Universidad São Paulo - USP, 2022, pp. 1–12.

[16] S. Clayman, A. Neto, F Verdi, S. Correa, S. Sampaio, I. Sakelariou, L. Mamatas, R. Pasquini, K. Cardoso, F. Tusa, C. Rothenberg, and J. Serrat. "The NECOS Approach to End-to-End Cloud-Network Slicing as a Service". In: IEEE Communications Magazine 59.3 (Mar. 2021).

[17] J. Nightingale, Qi. Wang, J. M. Alcaraz Calero, E. Chirivella-Perez, M. Ulbricht, J. A. Alonso-López, R. Preto, T. Batista, T. Teixeira, M. J. Barros, and C. Reinsch. "QoE-Driven, Energy-Aware Video Adaptation in 5G Networks: The SELFNET Self-Optimization Use Case". In: Int. J. Distributed Sens. Networks 12.1 (2016).

[18] P. Gouvas, A. Zafeiropoulos, C. Vassilakis, E. Fotopoulou, G. Tsiolis, R. Bruschi, R. Bolla, and F. Davoli. "Design, Development and Orchestration of 5G-Ready Applications over Sliced Programmable Infrastructure". In: 29th International Teletraffic Congress (ITC). Vol. 2. Sept. 2017, pp. 13–18.

[19] IETF. Framework for IETF Network Slices. RFC- Request for Comments draft-ietf-teas-ietf-network-slice framework-00. Internet Engineering Task Force, Mar. 2021, pp. 1–18.

[20] 3GPP. 5G-Evolution-3GPP. Technical Report Release 16-17. 3GPP, 2020, pp. 1–54.

[21] Telecommunication Standardization ITU-T. Framework of Network Virtualization for Future Networks. Tech. rep. ITU-T Y.3011. ITU-T - Telecommunication Standardization, Jan. 2012, pp. 1–28.

[22] ETSI. Mobile Edge Computing A key technology towards 5G. Technical Report WP No. 11. European Telecommunications Standards Institute, Sept. 2015.

[23] ONF - Open Networking Foundations. Applying SDN Architecture to 5G Slicing. Technical Report TR-526. ONF - Open Networking Foundations, 2016, pp. 1–19.

Mlungisi Molefe is a lecturer at the Durban university of Technology. He holds an MEng in electronic engineering from Durban University of Technology, his main research focus is network resource provisioning and optimisation for next generation wireless networks.

Khulekani Sibiya is an R&D Engineer at Dortech. He holds an MEng in Electronic engineering from Durban University of Technology, Main research focus is on cyber systems security.