Li, H., Assis, K., Vafeas, A., Yan, S., & Simeonidou, D. (2022). *Resilient and Energy Efficient DU-CU-MEC Deployments for Service Oriented Reliable Next Generation Metro Access Network*. Paper presented at 47th WWRF, Bristol, United Kingdom.

Resilient and Energy Efficient DU-CU-MEC Deployments for Service Oriented Reliable Next Generation Metro Access Network

Haiyuan Li, Karcius. D. Assis, Antonis Vafeas, Shuangyi Yan, and Dimitra Simeonidou

***Abstract*—Virtualization architecture with split baseband functions has received much attention from researchers and industries to provide a more flexible radio access network (RAN) to satisfy the new emerging diverse use cases including ultra-reliable low latency communication (uRLLC), enhanced mobile broadband (eMBB), and massive machine-type communications (mMTC). It divides functions aggregated in the active antenna unit and baseband unit into three more specific modules, including radio unit(RU), distributed unit (DU), and centralized unit (CU). Although this architecture helps vendors provide more flexible deployment selections in response to the diversity of NG network traffic, it raises a new problem regarding the highly timely manner of DU, CU, and MEC deployment management to realize on-time service, less power consumption and backup paths for higher reliability. In this paper, with the objective to minimize the power consumption of the next generation RAN architecture, we formulate a complete mixed integer linear programming (MILP) and a heuristic DCMH to optimize the baseband module deployment for both working and backup paths in the access network. The proposed solution is verified on a real deployed testbed in Bristol port. Results show that DCMH can realize similar performance as complete MILP. In addition, the longer idle time is, the more energy DCMH can save compared to keeping processing pools on standby. The project helps users get better services and operators save future RAN update costs.**

***Index Terms*—New generation RAN, DU-CU placement, energy efficiency, resilient protection, 5G and beyond**

I. INTRODUCTION

Nowadays, ever-increasing use cases such as ultra-reliable low latency communication (uRLLC), enhanced mobile broadband (eMBB) and massive machine-type communications (mMTC) place a severe burden on the current radio access network (RAN) architecture consisting of active antenna processing unit (AAU) and baseband processing unit (BBU) in terms of stringent constraints on latency and bandwidth of the fronthaul traffic [1][9][10].

The BBU and RRU are traditionally co-located in the same housing facility. Such a structure requires mobile operators to over-provision baseband resources to comply with the varying QoS, leading to high power consumption and poor scalability. Facing that, as shown in Figure 1, a new generation architecture has been proposed [2], that divides the physical layer and the date layer functions integrated into AAU and BBU into three-tier architecture, including radio unit (RU), distributed unit (DU) and core unit (CU). Fronthaul refers to the link between RU and DU, midhaul represents the link between DU and CU, and the last backhaul is the link between CU and UPF. The requests from user devices (UE) first arrive at the RUs distributed in cells, which are then integrated by RU and sent to MEC or 5G Core through DU, CU. The flexible deployment property of DU and CU allows mobile operators to provide diverse customized services to the users. In addition, functional virtualization further relieves the scaling pressure of base stations by implementing processing pools (PPs) which can accommodate both DU and CU and multi-access edge computing server (MEC).

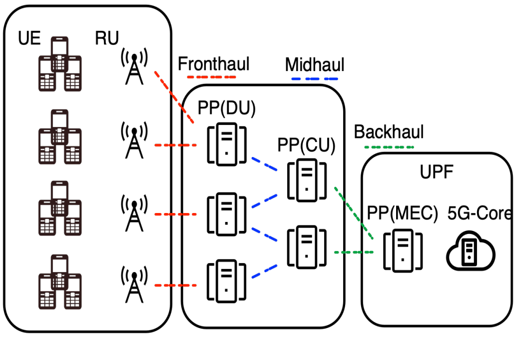


Figure 1: NG-RAN architecture

There are eight different functional splitting options as shown in Figure 2. Regarding to these options, as lower split between RU and DU left less physical layer functions in the RU, higher split solution require tighter fronthaul latency and wider bandwidth. In particular, we choose option 7 to split RU and DU, and option 2 to separate DU and CU [2]. As the supporting technology of the new generation RAN, compared to common public radio interface (CPRI), enhanced CPRI (eCPRI) highly reduces the fronthaul bandwidth and latency budget, increasing the feasibility of the DU distribution [3]. Furthermore, eCPRI can work on an open interface where the exclusive PP equipment can be shared between vendors and increase the industry convergence. In addition, since the function splitting introduces integrated fronthaul, midhaul and backhaul (X-Haul), to transmit the integrated traffic and meet the different requirements of multiple traffic, Flex-Ethernet interface was proposed to realize low latency switching between PPs [4].

In the context of NG-RAN, with diverse service requirements, operators must be able to react in a timely manner to decide on which PP to deploy the required DU and CU. However, this new architecture without an effective management solution might active a large number of PPs, resulting in excessive unnecessary energy consumption. Another notable issue in the current research is the network resiliency against PP or transmission path failure. Placing redundant resources is a solution that can increase network

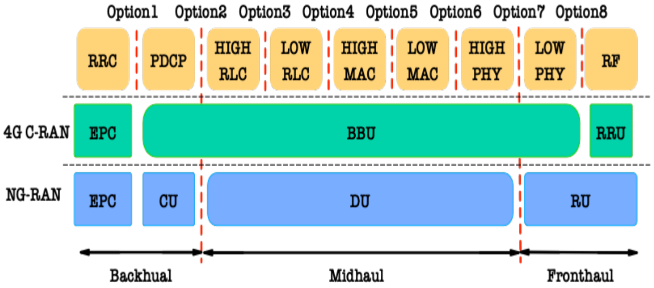


Figure 2: Function split options [2]

tolerance and service reliability at the expense of a high margin. In comparison, as discussed in [11], a backup solution DU, CU deployment decision with another routing path can also help to improve the reliability without margin pre-provision requirement. Therefore, in this paper, with the objective of minimizing the energy consumption, we designed a DU, CU placement policy with the 'best effort' resilience for the metro cellular network. The proposed deployment policy can meet the fronthaul delay, end to end delay, link bandwidth and other constraints while minimizing the energy consumption over the network. Besides, a complete mixed-integer linear programming problem is formulated and solved by CPLEX [5] as the optimal benchmark.

The reminder of this paper is organized as follows. Section II presents the next generation RAN scenario. Section III realizes the complete MILP formulation with formulated network constraints. To save the solving time, section IV proposes a heuristic solution. Section V provides some details of the testbed for our proposed solution. Section VI gives some results. Finally, section V conclude the paper.

II. SCENARIO

In the next generation RAN, RUs and PPs will be widely distributed through the cellular network. For easier of management of the network, in this paper, we divide the cellular network into clusters as shown in Figure 3. As PPs are not necessarily have the same resources, we divided it into PP1, PP2 and PP3. Among them, PP1 can only be used as DU; PP2 can be used as DU, CU and MEC with a small amount of resources for user plane functions (UPF) and PP3 can be used as DU, CU and MEC with more abundant UPF resources.

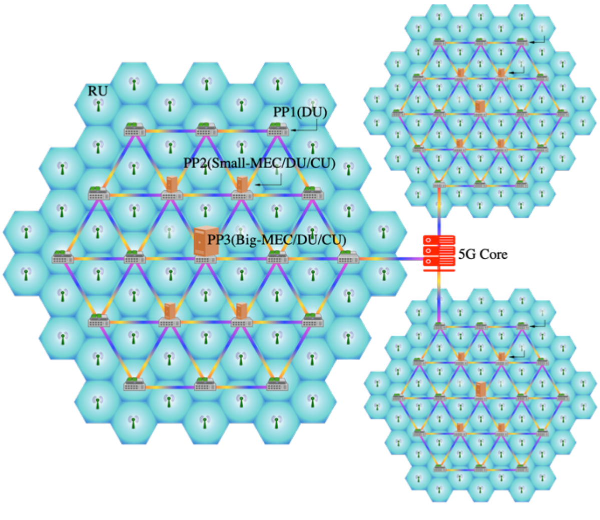


Figure 3: Cluster cellular network architecture

Table 1: Notations Used Throughout the Paper

	Notation	Description
Index	i and j	The source and destination nodes of the traffic demands in the network
	m and n	Endpoints of a physical link in the network
	f_i	It represents a function of type i where i=1 (DU), i =2 (CU), i=3 (MEC).
Given	$G=(N,E)$	A graph where N is the set of nodes and E is the set of edges (i.e. fiber links) connecting the nodes
	m_f	Function computing capability or maximum traffic rate (in Gbit/s)
	d_i	demand element, it is the requested traffic flow from node i(in Gbit/s)
	lat_{max}	Maximum end-to-end (ij) latency (in ms)
	$lat^{i,j}_{f,u}$	Maximum latency until node DU if f=1 (in ms)
	$dist_{mn}$	Distance between the nodes m and n on physical topology (in Km)
	χ	An arbitrarily large number
	ϵ	An arbitrarily small number
	$\rho_k(f,g)$	It is 1 (resp. -1) if f appears before (resp. after) g.
Variable	b_{ij}	Bandwidth of a virtual link which originates from node i and terminates at node j (in Gbit/s)
	q^{ij}_{mn}	The traffic that a demand from node i to node j uses in a fiber link m-n (integer).
	$x_{f,u}$	Number of f located on node u (integer).
	N_t	Total number of actives nodes in the network (integer).
	$y^{ij}_{f,u}$	It is 1 if function f handles traffic ij on node u (binary).
	$Y^{ij}_{f,u}$	It is 1 if ij meets its assigned f before or on node u (binary).
	T_u	Total number of function on node u.

For ease of reference, the notation used throughout the paper is summarized in Table 1.

III. MILP FORMULATION

In this section, we formulate a complete MILP to solve the problem of DU/CU/MEC deployment with resilience over a clustered PP network. We assume the power consumption of DU/CU and MEC are linear increasing with data size as [6][7][8]. In this vein, the power consumption of execution process will be fixed once the data size is acknowledged, and the overall energy consumption will not be affected by the deployment solution. The objective is to minimize the total energy in the network (in terms of activate nodes) to support all traffic demands with respective front and backhaul latency, where ε is chosen small enough so that the added term does not influence the main objective function

Objective function:

$$\min \sum_u p_u a_u + \epsilon N_t \tag{1}$$

Constraints:

Equation 2 is the traditional conservation constraints of flows. Equation 3 imposes there is traffic on physical link only if the demand i-j exist and passes on that link. Equation 3 expresses the fact that the traffic requested for a demand is the sum of the traffic flows established by source-destination pairs chosen.

$$\sum_j q^{ij}_{mn} - \sum_j q^{ij}_{nm} = \begin{cases} b_{ij} & m=i \\ -b_{ij} & m=j \\ 0 & else \end{cases} \tag{2}$$

$$q_{mn}^{ij} \le b_{ij} \qquad \forall i,j \tag{3}$$

$$\textstyle\sum_j b_{ij} \le d_i \qquad \forall ij \tag{4}$$

Equation 5 imposes an upper bound on bandwidth value of a link.

$$\textstyle\sum_{ij} q_{ij}^{mn} b_{ij} \le bw_{mn} \tag{5}$$

Equation 6 and Equation7 impose an upper bound on latency for backhaul and fronthaul, respectively.

$$\textstyle\sum_{mn} q_{ij}^{mn} dist_{mn} \le lat_{\{max\}} \tag{6}$$

$$\textstyle\sum_{mn} q_{mn}^{ij}\, y_{fu}^{ij} dist_{mn} \le lat_{fu}^{ij} \tag{7}$$

The following constraints state that a node cannot support more PPUs than its number of available cores 8, and that each function instantiation has a limited capacity and needs to be duplicated enough to handle all commodities that are assigned to it on the corresponding node 9:

$$x_{fu} \le cap_{fu} \tag{8}$$

$$\textstyle\sum_{ij} y_{fu}^{ij} b_{ij} d_i \le m_f x_{fu} \tag{9}$$

Equations 10, 11 and 12 define the total number of functions on a node and if a node is actived.

$$\textstyle\sum_f x_{fu} = T_u \tag{10}$$

$$\frac{T_u}{\chi} = a_u \tag{11}$$

$$\textstyle\sum_u a_u = N_t \tag{12}$$

Constraints 13, 14 and 15 state that at the source, a traffic has met none of the functions, whereas at the destination, it has met all the functions included in its chain, respectively.

$$Y_{fi}^{ij} = 0 \; \forall i,j,f \tag{13}$$

$$Y_{fj}^{ij} = 1 \; \forall i,j,f \tag{14}$$

$$\left(q_{mn}^{ij} - 1\right) + Y_{fn}^{ij} - Y_{fm}^{ij} \le y_{yn}^{ij} \; \forall i,j,m,n,f,g \tag{15}$$

Constraint 16 accounts for the order of the functions within the same chain: when a function g appears before f on the same chain. (Therefore, $\rho_k(f,g) = -\rho_k(g,f) = 1$).

$$Y_{fn}^{ij} - Y_{gm}^{ij} \le \rho_k(f,g) \; \forall i,j,m,n,f,g \tag{16}$$

IV. HEURISTIC

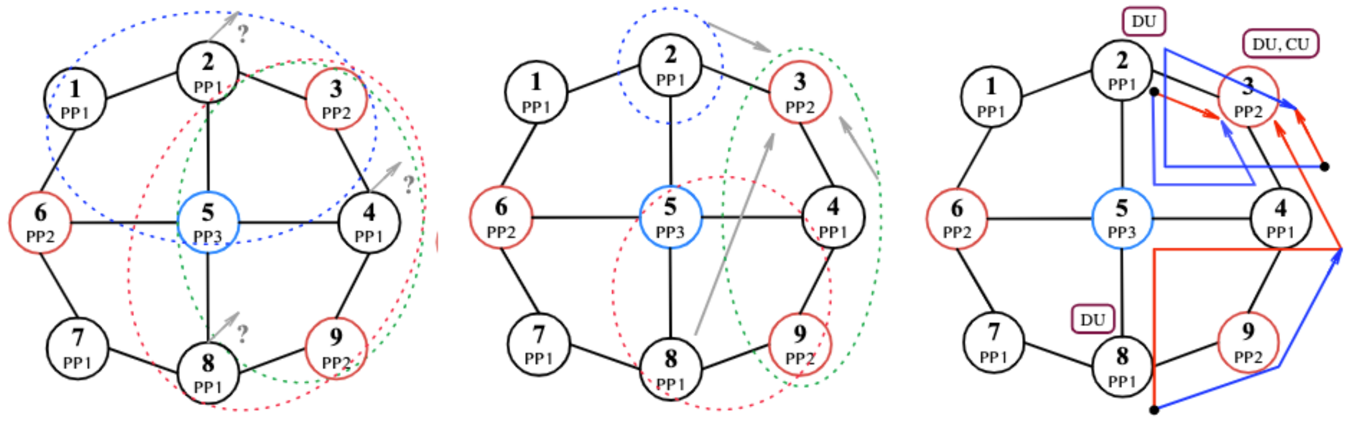
It is worth to notice that MILP is a time consuming and unscalable method that solves this NP-hard problem by exhaustive search in the solving space. It is not applicable to our problem when the network is large. Therefore, in this paper, an efficient heuristic (DCMH) is designed as an alternative to deployment both working and resiliency DU, CU and MEC with the objective of minimizing the system power consumption.

The heuristic realizes the objective function by 2 stages. In the first stage, based on the computing resource limitation of different PPs, the end to end latency and computing resource requirements of the traffic, DCMH decides whether traffic should be directed to 5G core or MEC, also how many MECs are required and where to deploy them. Based on the deployed MEC or 5G core as the destination of traffic, in stage 2, DCMH helps to find the working DU, CU and a backup solution under the constrains of fronthaul delay, bandwidth limitation and DU/CU resources. To be specific, DCMH is discussed in detail as follows.

In stage 1, for requests that can accept the end-to-end delay to the 5G core, their UPF will be deployed on 5G core to first save the limited resource on the MEC and second reduce the power consumption in the access network. Then, DCMH finds the requests that can accept the end-to-end delay to the pre-activated PP2 and PP3 and deploy MEC for them on these PPs to reduce the activation power consumption. As the resources of these pre-activated PPs are limited, to maximize the utilization, the decision order of these requests is sorted based on the occurrence times and less frequency, more priority. For instance, in Figure 4, Node 5 and Node 6 are pre-activated. If [Node 7, Node 8, Node 1] can get access to Node 6 and [Node 7, Node 8] can get access to Node 5, DCMH will make the decisions for requests to Node 6 in the order of [Node 1, Node 7, Node 8] in case the requests from Node 7 and Node 8 run out the resource of Node 6 causing a new PP activated for Node1, whereas Node 6 and Node 5 are supposed to support UPF for all these three requests. After that, DCMH sorted out the acceptable PP2s and PP3s for the rest traffic that doesn't have UPF. Among these PP2s and PP3s, who can be accessed by more traffic, the higher the priority to deploy MEC on it. In addition, under the assumption that the activation cost of PP3 is higher but less than twice that of PP2, DCMH puts PP3 before PP2 when choosing MEC as it is more affordable and can avoid activating two PP2s. In the end of Stage 1, as to provide protection path for resiliency, DCMH will figure out all pairs of paths for every traffic that are not sharing any intermediate node or paths on its way to UPF.

In stage 2, DCMH first deploy DU on the PPs that has already been activated before and after stage 1 with the consideration of their DU remaining resources. After that, DCMH applies the same method in stage 1 to sort the applicable PPs for deploying DU but changes the end-to-end latency constraint to stricter fronthaul requirements. While select the working and resiliency paths, as to minimizing the switch power consumption, DCMH chooses the pair with lowest hop summation, meanwhile, considers all bandwidth, DU, CU and MEC resource limitations.

To better explain the DCMH, Figure 4 provides a simple example of how DCMH deploys DU, CU and MEC in the



(a) Stage 1: MEC Deployment (b) Stage 2: DU deployment (c) Stage 2: Resilience and routing

Figure 4: Heuristic example explanation; In (a), (b), blue, green and red dash circles represent the farthest acceptable distance under end-to-end and fronthaul delay constraints of traffic from Node2, Node4 and Node8, respectively; In (c), red, blue arrows represents the working and resiliency paths.

RAN. In this network, there are 3 requests come from Node2, Node 4 and Node 8. In Figure 4a, the blue, green and red circles that correspond to 3 requests represent the farthest acceptable distance under the end-to-end delay limitation. Since only PP2 or PP3 can be activated as MEC, Node2 can get UPF service from Node 3 or Node 5; Node 4 and Node 8 can get UPF service from Node 3, Node 5 and Node 9. Taking the consideration of request and MEC computing resource of the PP2 and PP3, MEC may only be deployed on Node 3 to save the activation energy. In Figure 4b and 4c, three requests take Node 3 as destination. With the same definition, the circles in 4b now represent the fronthaul delay limitations. Traffic from Node 2 require a DU on itself as it cannot accept any other PP regarding to the fronthaul latency. DCMH deploys DU for Node 4 on Node 3 as it has already been activated when select CU. There will also be a DU deployed for traffic from Node8 since although Node8 can use DU on Node 5, Node 9 and itself, Node 5 and Node 9 are PP3 and PP2 and cost more power than Node 5. The last Figure 4c shows the working path in red and back up in blue. Under the end-to-end delay constrains, Node 8 cannot reach Node 3 without going through Node 4. Therefore, there is a tradeoff

Algorithm 1 DU/CU/MEC Deployment Heuristic

```
1: Collect the traffic (including fronthaul delay $T_f$, end to end delay $T_e$, computing resource requirement $R_m$, DU/CU resource requirement $R_d$ & $R_c$)
2: **Stage 1**: Deploy the UPF
3: for all traffic $i$ do
4:     if shortest path to 5G core < $T_{ei}$ then
5:         deploy 5G Core as UPF of $i$
6:         Find paths $\mathcal{P}_c < T_{ei}$
7:         Find the disjoint paths $\mathcal{P}_d$ from $\mathcal{P}_c$
8:     end if
9: end for
10: for traffic $i$ only accept MEC do
11:     for already been activated PP2 or PP3 $n_m$ do
12:         if i from $n_m$ then
13:             deploy MEC for $i$ on $n_m$
14:             Resource of $n_m$ - $R_{mi}$
15:         else
16:             if shortest path to $n_m$ < $T_{ei}$ then
17:                 store $n_m$ and i into $\mathcal{S}_0$
18:             end if
19:         end if
20:     end for
21: end for
22: count the emerging times of traffic in $\mathcal{S}_0$
23: Sort traffic in $\mathcal{S}_0$ according to occurrence frequency. Less frequency, more priority.
24: for Traffic $i$ in sorted $\mathcal{S}_0$ do
25:     for already been activated PP2 or PP3 $n_m$ do
26:         if Rest resource of $n_m$ > $R_{mi}$ then
27:             deploy MEC for $i$ on $n_m$
28:             Resource of $n_m$ - $R_{mi}$
29:         end if
30:     end for
31: end for
32: for rest traffic $i$ without UPF do
33:     Based on $T_{ei}$, store the acceptable PP2/PP3 into $\mathcal{S}_1$
34: end for
35: Sort PP2/3 in $\mathcal{S}_1$ according to occurrence frequency. More frequency, more priority.
36: for Node $n_m$ in $\mathcal{S}_1$ do
37:     for rest traffic $i$ without UPF do
38:         store $n_m$ and i into $\mathcal{S}_1$
39:     end for
40: end for
41: for PP2/PP3 $n_m$ in $\mathcal{S}_1$ do
42:     if $n_m$ is PP3 and $R_m$ of all corresponding traffic > Resource size of PP2 then
43:         Replace it in the first place
44:     end if
45: end for
46: count the emerging times of traffic in $\mathcal{S}_1$
47: Sort traffic in $\mathcal{S}_1$ according to occurrence frequency. Less frequency, more priority.
48: for traffic in $\mathcal{S}_1$ do
49:     if $n_m$ = origin of $i$ then
50:         Replace it in the first place
51:     end if
52: end for
53: repeat line 24-31 for $\mathcal{S}_1$
54: Find paths from origins to corresponding destinations satisfy the delay requirement
55: find all disjoint path pairs for resilience and store into $\mathcal{D}_0$
56: **Stage 2**: Deploy the DU/CU
57: for already been activated PP2 or PP3 $n_m$ do
58:     deploy DU/CU for traffic from it
59:     DU/CU resource - $R_d/R_c$
60: end for
61: for Already been activated PP1 $n_d$ do
62:     deploy DU for traffic from it
63:     DU resource - $R_d$
64: end for
65: for rest traffic $i$ without DU do
66:     based on $T_f$, store $i$, acceptable PPs and disjoint path pairs into $\mathcal{S}_2$
67: end for
68: for traffic $i$ in $\mathcal{S}_2$ do
69:     if only one PP1 choice and it is the origin of $i$ then
70:         deploy DU on this PP1
71:         DU resource - $R_d$
72:     end if
73: end for
74: for Traffic $i$ in $\mathcal{S}_2$ do
75:     Sorting the its disjoint path pairs based on the hop sum
76: end for
```

```
77: for rest traffic $i$ in $\mathcal{S}_2$ without DU do
78:     for all activated PP1/PP2/PP3 $n_a$ do
79:         for path pairs in $\mathcal{S}_2$ do
80:             if left DU resource of $n_a$ > $R_d$ and
                bandwidth of each hop on the path before DU > $R_d$
                and bandwidth of each hop after DU > $R_c$ then
81:                 Deploy DU on $n_a$
82:                 DU resource of $n_a$ - $R_d$
83:                 Store the pairs into $\mathcal{D}_1$
84:                 Use the shorter with $n_a$ path as the working path
                    $\mathcal{D}_2$, rest as the resilience $\mathcal{D}_3$
85:             end if
86:         end for
87:     end for
88: end for
89: for rest traffic $i$ without DU do
90:     store $i$ and its accessible DU into $\mathcal{S}_3$
91: end for
92: Sorting accessible DU according to occurrence frequency. PP1 first, then
    PP2. More frequency, more priority.
93: repeat line 77-88 for $\mathcal{S}_3$, store the pairs into $\mathcal{D}_1$, working path into $\mathcal{D}_2$
94: Pick one traffic, deploy CU on its destination, store $i$ and CU into $\mathcal{Q}$
95: for rest traffic $i$ on $\mathcal{D}_2$ do
96:     if intermediate node $n_t$ from DU to UPF of $i$ in $\mathcal{Q}$ and CU Resource
    of $n_t \geq R_{ci}$ then
97:         deploy CU for $i$ on $n_t$
98:         CU Resource of $n_t$ - $R_{ci}$
99:     else
100:         deploy CU for $i$ on its destination
101:         CU Resource of destination of $i$ - $R_{ci}$
102:     end if
103: end for
104: Repeat line 95-103 for backup path $\mathcal{D}_1$
105: if left traffic without DU/CU/MEC caused by bandwidth or resource
     limitation then
106:     applying first fit or reject
107: end if
```

between resilience and our objective. The optimal power consumption of DCMH is realized at the expense of only providing 'best effort' protection.

V. TESTBED

To verify the effectiveness of our proposed solution on baseband module distribution, based on the same function splitting option 2 and 7, we build am NG-RAN testbed in Bristol port. As show in Figure 8, there are 8 nodes corresponding to the PP2 and PP3 in our simulation. Their virtual topology is shown in the left top gray diagram. The figure with blue dash lines represents the IoT sensors and 5G gateway that collects information through Wifi and send the packaged information to PP. Two different PPs, PP2 and PP3 are shown in green and yellow dash lines.

VI. RESULTS

With the objective to show the performance of our proposed solution, taking complete MILP as the optimal benchmark, we evaluate the performance of DCMH, all awaken (keep all server awake), first fit (active based the resource when there are requests) [12], and random allocation (activate randomly with the premise of satisfying the request) in a small network as shown in Figure 6.

With the assumption that PP number 3 and 5 have been pre-activated and all seven PP receive requests over a period

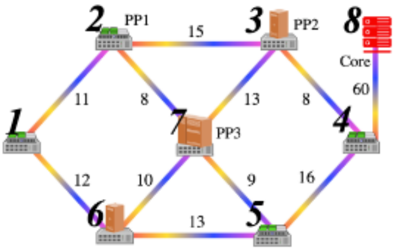


Figure 6: Small network example (united in km)

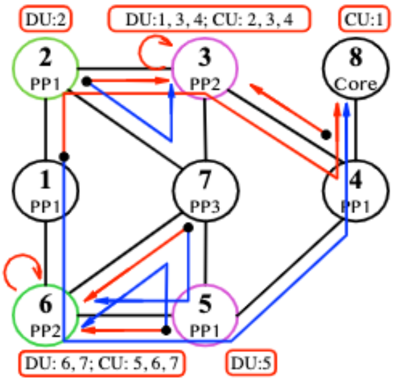


Figure 7: DU, CU deployment for small network as shown in Figure 6

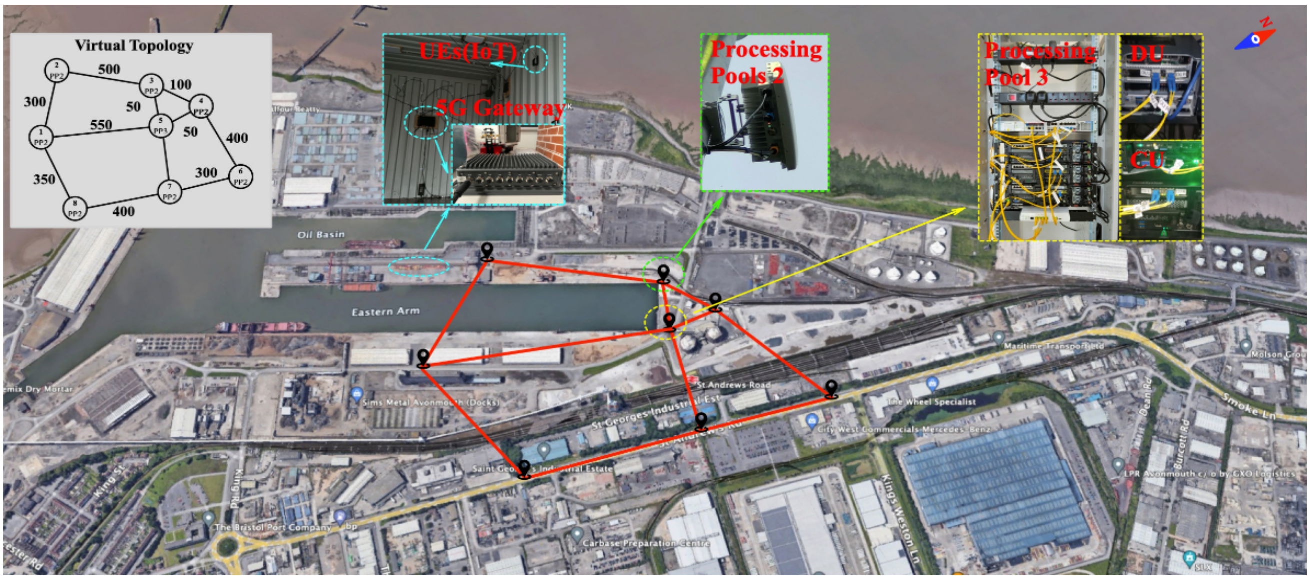


Figure 8: The testbed at Bristol port. The gray graph represents the virtual topology of the network. Blue dash lines include the UE related devices. Yellow and blue dash lines include the PP3 and PP2 infrastructures. (Figure is take from Google Earth, earth.google.com)

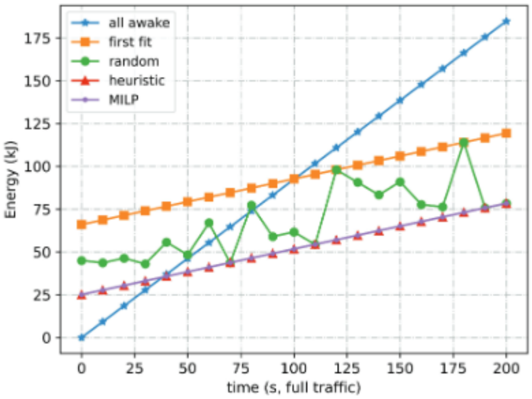


Figure 9: Power consumption of 7 requests with PP 3 and 5 preactivated

of time, Figure 7 shows the details of DU/CU/MEC deployment of DCMH. Green circle means new activated PPs, pink circles means the PPs that have already been activated; Red arrow represents working path, blue one represent resilience; DU, CU deployment is marked on the node where the number indicates the source PP. Figure 9 provides the energy consumption of different management solutions over different waiting times in the unit of kJ. All standby modes waste more energy as the wait time increases. In contrast, if no request is received for more than 37s, DCMH is a better solution.

VII. CONCLUSION

Facing the diversity of next generation traffic, baseband function splitting will become an effective solution to help operators to provide users better customized service. However, this solution also brings higher energy cost as more facilities will be implemented. To decrease the energy consumption and increase the economic benefits, our designed DCMH network management method can dramatically outperform all awaken mode and other solutions. It improves the achievability of the next generation RANs.

ACKNOWLEDGMENT

The authors like to thank the support from 5G O-RAN, China Scholarship Council and European Commission's Horizon 2020 research and innovation program under grant agreement No 871428, 5G-CLARITY project.

REFERENCES

[1] Klinkowski M. Latency-Aware DU/CU Placement in Convergent Packet-Based 5G Fronthaul Transport Networks[J]. *Applied Sciences*, 2020, 10(21): 7429..

[2] 3GPP, "Radio access architecture and interfaces (release 14)," TR 38.801, 2017..

[3] Perez G O, Lopez D L, Hernandez J A. 5G new radio fronthaul network design for eCPRI-IEEE 802.1 CM and extreme latency percentiles[J]. *IEEE Access*, 2019, 7: 82218-82230.

[4] Li R, Wang R, Halachmi N, et al. X-Ethemet: Enabling integrated fronthaul/backhaul architecture in 5G networks[C]//*2017 IEEE Conference on Standards for Communications and Networking (CSCN)*. IEEE, 2017: 121-125..

[5] LOG, Inc, "ILOG CPLEX: High-performance software for mathematical programming and optimization," 2018, see https://www.ibm.com/analyti

[6] Xiao Y, Zhang J, Ji Y. Energy-efficient DU-CU deployment and lightpath provisioning for service-oriented 5G metro access/aggregation networks[J]. *Journal of Lightwave Technology*, 2021, 39(17): 5347-5361.

[7] Zorello L M M, Troia S, Quagliotti M, et al. Power-aware optimization of baseband-function placement in cloud radio access networks[C]//*2020 International Conference on Optical Network Design and Modeling (ONDM)*. IEEE, 2020: 1-6.

[8] Zorello L M M, Sodano M, Troia S, et al. Power-efficient baseband-function placement in latency-constrained 5G metro access[J]. *IEEE Transactions on Green Communications and Networking*, 2022.

[9] Harutyunyan D, Shahriar N, Boutaba R, et al. Latency-aware service function chain placement in 5G mobile networks[C]//*2019 IEEE Conference on Network Softwarization (NetSoft)*. IEEE, 2019: 133-141.

[10] Yu H, Musumeci F, Zhang J, et al. DU/CU Placement for C-RAN over optical metro-aggregation networks[C]//*International IFIP Conference on Optical Network Design and Modeling. Springer*, Cham, 2019: 82-93.

[11] Masood Khorsandi B, Colle D, Tavernier W, et al. Dedicated and shared solutions for reliable function splitting in 5G X-haul[J]. *Photonic Network Communications*, 2020, 40(3): 245-255.

[12] Bays C. A comparison of next-fit, first-fit, and best-fit[J]. *Communications of the ACM*, 1977, 20(3): 191-192.

Haiyuan Li received the B.Sc degree in communication engineering in Central South University, China, in 2019, and the M.Sc degree in communication networks and signal processing from University of Bristol, U.K., in 2020. He is currently pursuing his Ph.D. degree with the school of electrical and electronic engineering in University of Bristol. His research interests include Mobile Edge Computing, Machine Learning, Game Theory and Network Optimization.

Karcius Day R. Assis received the B.Sc. degree in electrical engineering from the Federal University of Paraíba (currently Federal University of Campina Grande) in 1998, the M.Sc. degree from the Federal University of Espírito Santo in 2000, and the Ph.D. degree in electrical engineering from the University of Campinas in 2004. From 2004 to 2010, he was with Salvador University, Federal University of ABC-UFABC and Federal University of Recôncavo da Bahia, Brazil. In 2010, he joined the Department of Electrical and Computer Engineering, Federal University of Bahia, Salvador, Brazil, where he is an Associate Professor. Currently, Karcius has returned to the High Performance Networks Group, Smart Internet Lab at the University of Bristol, U.K., as a visiting researcher, having held this position in 2015. His research interests include Optical Networking, Traffic Engineering, Resource Management, Network Survivability, Machine Learning, Future Internet, and Telecommunication Networks Optimization and Projects. He won an International Visiting Fellowships in 2017 and spent some time at Essex University, United Kingdom.

Dr Antonis is a Senior Research Associate in 5G Network Systems Engineering & Administration at University of Bristol. He specialises in multimodal sensing, wireless networking and embedded hardware

Dr Shuangyi Yan is a Senior Lecturer in the High Performance Networks Group in the Smart Internet Lab at the University of Bristol. He received the B.E degree in information engineering from Tianjin University, Tianjin, China in 2004. In 2009, he got the PhD degree in Optical Engineering from Xi'an Institute of Optics and Precision Mechanics, CAS, Xi'an, China. From 2011 to 2013, Dr Yan worked in the Hong Kong Polytechnic University, Hong Kong, as a postdoctoral researcher, investigating on the spectra-efficient long-haul optical transmission system and low-cost short-range transmission system. In July 2013, he joined the University of Bristol. His research focuses on machine-learning applications in dynamic optical networks and 5G Beyond networks, programmable optical networks, and data centre networks. He is the author or co-author of over 80 refereed publications, of which include several post-deadline papers in optical communication related top-level conferences. Dr Yan has delivered 9 invited talks in relevant conferences, such as Optical Networking and Communication Conference (OFC 2019), Optical Network Design and Modelling (ONDM 2019 \& 2018), ECOC 2018, and Asia Communications and Photonics Conference (ACP 2017). He also served as Technical Program Committee members and Co-Chairs in several conferences, such as EuCNC, ONDM, OECC, and ACP.

Dimitra Simeonidou (FREng, FIEEE) is a Full Professor at the University of Bristol, the Co-Director of the Bristol Digital Futures Institute, and the Director of the Smart Internet Lab. Her research focuses in the fields of high performance networks, programmable networks, wireless-optical convergence, 5G/B5G and smart city infrastructures. She is increasingly working with Social Sciences on topics of digital transformation for society and businesses. Dimitra has also been the Technical Architect and the CTO of the Smart City project Bristol Is Open. She is currently leading the Bristol City/Region 5G urban pilots. She is the author and co-author of over 500 publications, numerous patents and several major contributions to standards. Dimitra is a Fellow of the Royal Academy of Engineering, a Fellow of the Institute of Electrical and Electronic Engineers, and a Royal Society Wolfson Scholar.