This is a postprint version of the following published document: Ruiz Pérez, Lidia; Durán Barroso, Ramón José; Miguel Jiménez, Ignacio de; Merayo Álvarez, Noemí; Aguado Manzano, Juan Carlos; Fernández Reguero, Patricia; Lorenzo Toledo, Rubén Mateo; Abril Domingo, Evaristo José. Joint VNF-Provisioning and Virtual Topology Design in 5G Optical Metro Networks. In: 2019 21st International Conference on Transparent Optical Networks (ICTON). Angers, France: IEEE, 2019

Joint VNF-Provisioning and Virtual Topology Design in 5G Optical Metro Networks

L. Ruiz, R.J. Durán, I. de Miguel, N. Merayo, J.C. Aguado, P. Fernández, R.M. Lorenzo and E.J. Abril
Optical Communications Group. Universidad de Valladolid. Paseo de Belén, 15, 47011, Valladolid, Spain
e-mail: lruiper@ribera.tel.uva.es, rduran@tel.uva.es

ABSTRACT
5G technology will provide networks with high-bandwidth, low latency and multitenancy. The integration of computing and storage resources in the edge of the fronthaul network, i.e., multi-access edge computing (MEC), will allow to instantiate some virtual network functions (VNF) in those computing resources. The backhaul of 5G networks will be based on optical technology, in particular WDM, due to its high capacity and flexibility. In this paper, we analyse the problem of VNF-provisioning in a metro ring-topology network equipped with MEC resources and with a WDM network connecting the edge nodes. In contrast to previous proposals, the method decides where VNFs must be instantiated but also the design of the virtual topology for the WDM metro network in order to reduce the service blocking ratio and the number of resources in operation.

Keywords: 5G, MEC, VNF-Provisioning, VNF-Placement, VNF-Chaining, virtual topology design.

1. INTRODUCTION

Some of the objectives of 5G are to offer multi-tenancy and high-capacity communications while reducing the latency, among other new features. In order to achieve those key performance indicators, 5G networks require profound architectural changes compared to previous mobile communication architectures. In the new architecture, different computing and storage resources will be deployed along the network infrastructure allowing not only the cloud computing paradigm but also multi-access edge computing (MEC) and fog computing. Thanks to the latter proposals, the distance between users and the computing resources can be reduced and, therefore, the latency will also decrease fulfilling the requirements of the new verticals.

Two key technologies for that evolution in 5G networks equipped with MEC resources are software defined networking (SDN) and network function virtualization (NFV). SDN disaggregates the control and the data plane and facilitates network management. NFV is a networking paradigm for which network functions (NF) are deployed as software instances, called virtual network functions (VNF), hosted in commodity servers, instead of the traditional, vendor-based hardware appliances. These servers can be located at data centres but they can also be located at the edge nodes of the networks thanks to the MEC paradigm, which uses distributed computing techniques to foster collaboration between them. In this way, VNFs can be placed closer to the end-users, allowing for the achievement of low-latency requirements.

When instantiating VNFs, operators must decide both the number of VNFs and their location. They also need to consider that, in order to deploy a network service, its associated traffic must traverse a number of these VNFs in a given order, creating the so-called service chain (SC). Therefore, operators must also select the instantiated VNFs that will compose the SC. These two problems are known as the VNF-placement and chaining problems.

In [1] authors present an ILP that optimally places the VNFs for serving end-to-end requests. Authors in [2], [3] study the impact of latency in the VNF location. Latency is also tackled in [4], where Askari et al. propose an algorithm for dynamic service chaining in a metro-area network, which minimizes the VNF-enabled nodes and the blocking probability and ensures that the latency requirements are met. In [5], [6], authors propose a heuristic algorithm to solve the chaining problem in a realistic scenario, with drastic changes of expected traffic, with the aim of minimizing the service blocking rate.

In [7], we present a genetic algorithm for effective service mapping (GASM) that solves the VNF placement problem in a 5G metro network with a star topology, like the one presented in [5],[6]. In our proposal, different edge nodes (each one equipped with MEC resources) are directly connected to a central office, CO (equipped with computing cloud resources). Then, the VNFs can be instantiated in MEC nodes or in CO but there is not collaboration between MEC resources, i.e., the SC for a certain service started at an edge node can only use the VNFs instantiated in its local MEC resources or those VNFs instantiated in CO computing resources. That algorithm aims at minimizing the service blocking ratio and the computing resource consumption, and considers both the computing and bandwidth resource availability to decide the number and location of the VNFs.

In this study, we extend that work by solving the VNF-placement and chaining problems in a ring topology WDM metro network, and analysing the benefits of the collaboration of MEC nodes to solve each service request. The network consists of a CO equipped with cloud computing resources and 5G edge nodes equipped with MEC resources. If two consecutive VNFs in a SC are instantiated in different nodes, a virtual link must be created between those nodes to accommodate the capacity required by the SC. These virtual links can be electrically groomed in existing lightpaths with enough capacity, or lead to the creation of one or more new lightpaths between those two nodes. Therefore, the algorithm must not only solve the VNF-placement and chaining problems but also the virtual topology design and embedding problem, which consists of the following three subproblems:

- Connectivity problem: which lightpaths must be established, i.e., the virtual topology
- The routing and wavelength assignment problem (RWA) for each lightpath.
- The grooming of the virtual links in the lightpaths of the virtual topology.

2. RING METRO NETWORK EQUIPPED WITH MEC RESOURCES

As stated in the introduction, we assume that 5G nodes (equipped with MEC resources) and the CO (equipped with computing cloud resources) are connected in a WDM ring. We consider two kinds of 5G nodes: The High-Demand 5G Nodes (HD-5G-Ns) and the Low-Demand 5G Nodes (LD-5G-Ns). The difference between them are the maximum number of connected users and IT resources they are equipped with. Each bidirectional link in the network is composed of two fibres (one per each direction). Then, each node has a ROADM as well as optical transceivers. Fig. 1 shows the network architecture.

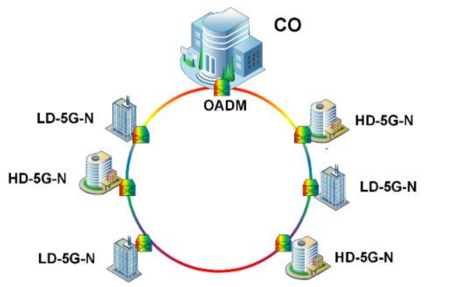

Figure 1. Ring metro 5G network with MEC.

3. A GENETIC ALGORITHM TO SOLVE THE VNF-PLACEMENT, CHAINING AND VIRTUAL TOPOLOGY DESIGN AND EMBEDDING PROBLEMS

The new algorithm, genetic algorithm for service mapping and virtual topology design and embedding (GASM-VTD), enhances GASM [7] by including the design of the virtual topology for the WDM ring metro network (besides solving VNF-placement and chaining problems). Each potential solution (VNF placement, SC for each service requested by the users and virtual topology) is represented by an individual. These individuals are described by a chromosome composed by genes. The chromosome of GASM-VTD is equal to that of GASM. Each gene encodes the number of instances of a given VNF that a specific node of the network host in that solution (see [7] for more details). However, the translation process is different. When translating the chromosome, GASM-VTD instantiates the VNFs according to the chromosome information. If there are not enough computing resources (CPU, memory or hard disk drive, HDD) to instantiate all the VNFs determined by the chromosome, that individual is discarded. If all VNFs have been instantiated, GASM-VTD sorts the service requests according to the operator's chosen priority and it tries to create the SC associated to each requested service. We devise two versions of GASM-VTD:

- GASM-VTD-No-Collaborative: only utilizes VNFs instantiated in local 5G-Ns or in CO. The order of the searching process is local 5G-N $\rightarrow$ CO for each VNF of the SC. That policy is the one described in [5]–[7] with the name "MEC-First". GASM-VTD-No-Collaborative will provide the same performance of GASM [7] when there are no restrictions in the capacity of the network.
- GASM-VTD-Collaborative: it can use any VNF instantiated in any node of the network. The order of the searching process is local 5G-N $\rightarrow$ CO $\rightarrow$ HD-5G-Ns $\rightarrow$ LD-5G-Ns.

In both versions, when two consecutive VNFs in the SC are placed in two different nodes, a virtual link between them with enough capacity for the service must be established. For that process, traffic grooming is allowed. Hence, if there is one or more lightpaths with enough capacity already established between those nodes, they will be used to set up that virtual link. Otherwise, a new lightpath will be established (if there are enough available optical resources). In order to deploy new lightpaths, GASM-VTD uses shortest path and first-fit [8] to solve the RWA problem. If there are not enough computing and/or network resources to establish the SC of a service request, that request is blocked (and therefore, the service blocking ratio increases). At the end of the process, GASM-VTD provides the VNF-Mapping, the SC for each service request and the virtual topology established in the metro optical network.

GASM-VTD follows the classical loop of genetic algorithms [9]. Initially, it generates an initial population of individuals (i.e., potential solutions). Similarly to [7], the initial population is randomly created except two ad-hoc individuals included to speed-up the searching process. The first one is built considering that the number of instances of VNF in each node (i.e., the gene value) is determined using "MEC-First" policy [5]–[7]. The second ad-hoc individual is built using "CO-First" [5]–[7]: the required VNFs are computed following the placement order CO $\rightarrow$ local 5G-N. Once the initial population is constructed, the individuals undergo genetic operations called crossover and mutation. In the crossover operation, GASM-VTD randomly selects two individuals of the set. Then, it also chooses a random crossover point and interchanges the second part of the chromosomes, creating two new individuals. In mutation, the GASM-VTD mutates each gene (i.e., randomly change the gene value) according to a user-defined mutation probability. In each generation GASM-VTD selects the fittest individuals among the parents and the offspring to be the parents of the following generation. The process is repeated until the number of generations reaches a value determined by the user.

The fitness of each individual is determined translating the chromosome and evaluating three parameters: service blocking ratio, number of CPU cores in operation and the number of wavelengths used, in that order. Therefore, the main parameter is the service blocking ratio. The number of CPU cores is used to solve the ties in the first parameter, and the number of wavelengths is used if there are still ties.

4. SIMULATION STUDY AND RESULTS

We conduct a simulation study in OMNeT++ (a C++ based discrete event simulator) to test the performance of GASM. We have selected a WDM ring topology connecting a CO, 5 HD-5G-Ns and 5 LD-5G-Ns. The architecture of the network is the one presented in Section 2. Each node is equipped with a set of transceivers and a ROADM to allow the use of 10 wavelengths channels at 10 Gb/s. We assume that the CO, the HD-5G-Ns and the LD-5G-Ns are all VNF-enabled and equipped with the IT resources shown in Table 1 [5], [6]. We assume that the network operator offers three kind of services, VoIP, video streaming and web searching. Each service has an associated SC, with the VNFs and bandwidth requirements shown in Table 2 [2], [3], [5]–[7]. Furthermore, each VNF has the IT resource requirements shown in Table 3.

Table 1. IT resource distribution in CO and AOs

Location	Computational resources
CO	100 CPU cores, 480 GB RAM and 27 TB HDD
HD-5G-Ns	16 CPU cores, 64 GB RAM and 10 TB HDD
LD-5G-Ns	8 CPU cores, 32 GB RAM and 7 TB HDD

Table 2. Requirements of the deployed service chains

Service	Chained VNFs*	Bandwidth
VoIP	NAT-FW-TM-FW-NAT	64 kbps
Video	NAT-FW-TM-VOC-IDPS	4 Mbps
Web Services	NAT-FW-TM-WOC-IDPS	100 kbps

* NAT:*Network Address Translator*, FW: *Firewall*, TM: *Traffic Monitor*, WOC: *WAN Optimization Controller*, VOC: *Video Optimization Controller*, IDPS: *Intrusion Detection Prevention System*.

Table 3. Hardware requirements associated to the VNFs.

Service	HW requirements.	Throughput	
NAT	CPU: 2 cores, RAM: 4 GB, HDD: 16 GB*	2 Gbps [10]	* We assume that FW also includes NAT function, therefore the HW requirements are the same. No datasheet for VOC was found, so the figures are derived from the figures of the other VNFs.
FW	CPU: 2 cores, RAM: 4 GB, HDD: 16 GB	2 Gbps [10]	
TM	CPU: 1 core, RAM: 2 GB, HDD: 16 GB	1 Gbps [11]	
VOC	CPU: 2 cores, RAM: 4 GB, HDD: 2 GB	2 Gbps*	
WOC	CPU: 1 core, RAM: 2 GB, HDD: 40 GB	0.5 Gbps [12]	
IDPS	CPU: 2 cores, RAM: 4 GB, HDD: 8 GB	1 Gbps [13]	

We include an input parameter, $\bar{u}$, which represents the average number of users per HD-5G-Ns. In each simulation, the number of connected end-users to each HD-5G-Ns is randomly selected using a uniform distribution between $[0, 2\bar{u}]$. Similarly, the average number of users connected to each LD-5G-Ns is randomly chosen using and uniform distribution between $[0, 2\bar{u}/10]$. Each user can request a service, either VoIP, video or web services, with a probability of 30%, 20% or 50% respectively [7]. We create an initial population composed by 5 individuals, as the starting point of the evolutive process. Furthermore, we create 50 generations of 10 individuals each. We repeat the simulation 500 times with different traffic demands. Results from the two versions of GASM-VTD are compared with those of the "MEC-First" and "CO-First" policies [5]–[7]. Fig. 2 shows the service blocking ratio while Fig. 3 shows the percentage of active CPU cores depending on the average number of users in HD-5G-N, $\bar{u}$. Both figures are plotted with 95% confidence intervals.

Results show that GASM-VTD-No-Collaborative obtains the same service blocking ratio than MEC-First (Fig. 2) but using a lower number of active CPU cores and, therefore, less energy consumption (Fig. 3). Moreover, the solutions provided by GASM-VTD-No-Collaborative are better than those provided by CO-First (more than two order of magnitude) in terms of service blocking ratio which is the main parameter to be optimized.

However, the main advantage of the architecture and the algorithm presented in this paper is when GASM-VTD allows the collaboration of all 5G nodes (equipped with MEC) to solve the VNF-placement and chaining problems. In those conditions, the solutions provided by GASM-VTD-Collaborative lead to even lower blocking service ratio than those of the version without collaboration (Fig. 2). In fact, the system (network and MEC resources) supports, in average, 1,000 and 100 additional users in each HD-5G-N and LD-5G-N, respectively, without increasing the service blocking ratio in Fig. 2 (note that the average number of LD-5G-N users is 10% of the average users of HD-5G-N).

Regarding the number of active CPUs (Fig. 3), the collaborative version utilizes the same number of CPU cores than the non-collaborative one for an average number of users in each HD-5G-N lower than 5,000 (i.e., in the region where service blocking ratio is under 10^{-2}) and therefore, GASM-VTD-Collaborative reduces the service blocking ratio without increasing the energy consumption of the network. For higher number of users, GASM-VTD-Collaborative utilizes more CPU cores than the other methods in order to reduce the service blocking ratio.

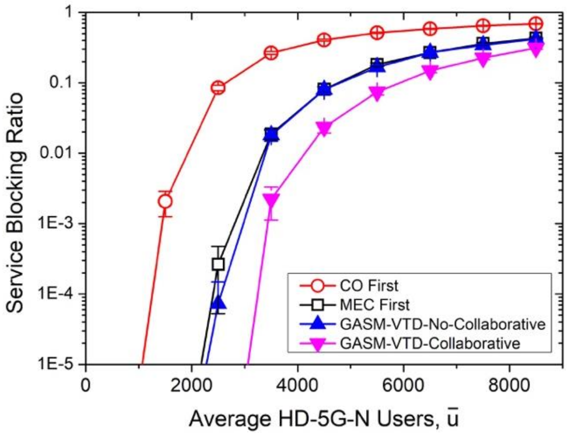

Figure 2. Service blocking ratio.

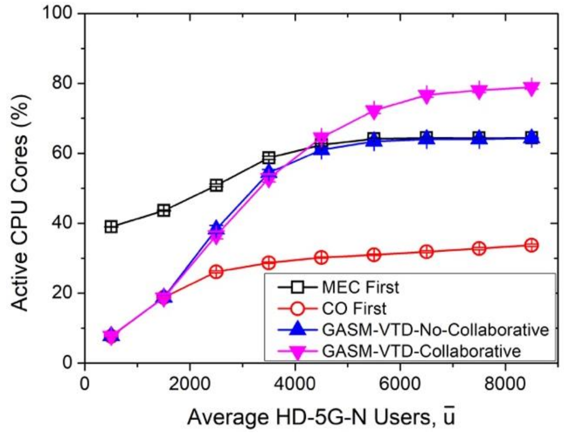

Figure 3. Percentage of active CPU cores.

5. CONCLUSIONS

In this paper we have evaluated the benefit of integrating different MEC resources in a 5G WDM metro network with ring topology allowing the collaboration of all MEC nodes to solve the service requests. A new genetic algorithm has been proposed to solve not only the VNF-placement and chaining problems but also the virtual topology design of the WDM network. Results show that thanks to the new proposal, the service blocking ratio can be reduced while optimizing the number of active CPU cores.

ACKNOWLEDGEMENTS

This work has been supported by Spanish Ministry of Economy and Competitiveness (TEC2017-84423-C3-1-P), the fellowship program of the Spanish Ministry of Industry, Trade and Tourism (BES-2015-074514).

REFERENCES

[1] T. Lin *et al.*: Optimal network function virtualization realizing end-to-end requests, in *Proc. 2015 IEEE Global Communications Conference (GLOBECOM)*, San Diego, CA, USA, 2015, pp. 1–6.

[2] M. Savi *et al.*: Impact of processing costs on service chain placement in network functions virtualization, in *Proc. of the 2015 IEEE Conference on Network Function Virtualization and Software Defined Network (NFV-SDN)*, San Francisco, CA, 2015, pp. 191–197.

[3] M. Savi *et al.*: To distribute or not to distribute? Impact of latency on Virtual Network Function distribution at the edge of FMC networks, in *Proc. 2016 18th International Conference on Transparent Optical Networks (ICTON)*, 2016, pp. 1–4.

[4] L. Askari *et al.*: Virtual-network-function placement for dynamic service chaining in metro-area networks, in *Proc. 2018 International Conference on Optical Network Design and Modeling (ONDM)*, Dublin, 2018, pp. 136–141.

[5] J.-J. Pedreno-Manresa *et al.*: Dynamic QoS/QoE assurance in realistic NFV-enabled 5G Access Networks, in *Proc. 2017 19th International Conference on Transparent Optical Networks (ICTON)*, Girona, Spain, 2017, pp. 1–4.

[6] J.-J. Pedreno-Manresa *et al.*: On the Need of Joint Bandwidth and NFV Resource Orchestration: A Realistic 5G Access Network Use Case," *IEEE Communications Letters*, vol. 22, no. 1, pp. 145–148, Jan. 2018.

[7] L. Ruiz *et al.:* A Genetic Algorithm for VNF Provisioning in NFV-Enabled Cloud/MEC RAN Architectures, *Applied Sciences*, vol. 8, no. 12, p. 2614, Dec. 2018.

[8] H. Zang *et al.*: A review of routing and wavelength assignment approaches for wavelength-routed optical WDM networks, *Optical Networks Magazine*, vol. 1, no. 1, pp. 47–60, 2000.

[9] D. Goldberg, "Genetic algorithms in optimization, search and machine learning," *Reading: Addison-Wesley*, 1989.

[10] Juniper Networks, vSRX Virtual Firewall, Data Sheet, Nov-2018.

[11] Brocade Communications System, Virtual Traffic Manager, Data Sheet, 2015.

[12] Talari Networks, Talari SD-WAN Solutions, Data Sheet, 2018.

[13] Cisco, Cisco Adaptive Security Virtual Appliance (ASAv), Data Sheet, 2018.