MEC Placement Problem in Protected 5G Networks

Hernâni D. Chantre[1,2], **Nelson L.S. da Fonseca**[1]

[1]Institute of Computing – State University of Campinas (UNICAMP)
Campinas – SP – Brazil

[2]Faculty of Science and Technology – University of Cabo Verde (UniCV)
Praia – Cabo Verde.

hernani@lrc.ic.unicamp.br, nfonseca@ic.unicamp.br

Abstract. *This paper summarizes the thesis "MEC Placement Problem in Protected 5G Networks", which tackles the question of where to position Multi-access edge computing nodes (MECs) in 5G networks to achieve stringent 5G requirements at a minimum cost. The thesis formulated the MEC placement problem with traditional protection schemes to provide redundancy in the event of MEC failures, and proposed a new scheme called 'enhanced shared protection* $1 : N : K$*'. The work investigated the impact on the costs of new variations of the MEC placement approaches considering different distributions of VNFs forming SFC requests. The results indicate that the enhanced protection scheme is cost-effective, and the VNF distribution impacts significantly on the cost.*

1. Introduction

The fifth-generation cellular network (5G) supports massive broadband applications, massive machine-type communication, and large-scale connectivity for the Internet of Things (IoT). As mobile data traffic continues to increase, there is a need for mechanisms that can improve network resilience and optimize the use of network resources. To support reliable Service Function Chains (SFC) requests, 5G network resources are divided into logical resources or slices using cloud-slicing. Multi-access Edge Computing nodes (MECs) leverage virtualization, thus, providing cloud slices to host VNF chains [Medeiros et al. 2023]. MECs usually have limited computational and storage capacities as well as bandwidth to handle the traffic generated by diverse service requests. Moreover, VNFs, slices, VMs, and MECs are prone to failures and malfunctions, which can compromise the reliability required for 5G services. Hence, achieving high levels of reliability in 5G networks with MECs is challenging.

This paper summarizes the thesis "MEC Placement Problem in Protected 5G Networks" [Chantre and da Fonseca 2023] that attempted to answer the following question "Where to position MECs in 5G networks to achieve 5G reliability and latency requirements at a minimum cost". The work in the thesis considered a set of demand points requesting the execution of SFCs with stringent requirements in terms of latency and reliability and a set of locations that can host an MEC. The MEC placement problem in 5G network aims at devising locations to MECs and hosted slices so that the requirements of services can be supported and cost minimized.

The Infrastructure Provider (InP) must choose a protection scheme to support the 5G reliability and latency requirements at a minimum cost. Unlike existing works in the literature that consider reliability constraints, this thesis evaluated different protection schemes and levels of reliability constraints on the problem formulation. Specifically, traditional protection schemes such as dedicated $1:1$ protection and shared $1:N$ protection were investigated to provide redundancy in the event of MEC failures, unlike existing literature that only considers reliability constraints. In the dedicated protection scheme $1:1$, the backup (secondary) slice is reserved exclusively for one demand point. The primary and the backup slice of a demand point cannot be hosted on the same MEC node. In the shared protection scheme $1:N$, one primary slice is allocated to a single demand point, whereas the secondary slice is shared by multiple $N < |V|$ demand points.

The thesis proposed a new protection scheme called 'enhanced shared protection $1:N:K$' that offers additional redundancy levels. The proposed enhanced shared protection scheme notated as $1:N:K$, consists of extending the shared protection $1:N$ with $K-2$ additional redundant slices to mitigate the impact of failures of MEC nodes. These $K-2$ redundant nodes can be shared by all the demand points. Thus, if the secondary slice fails, the demand point is still protected by additional $K-2$ redundant MECs nodes.

Moreover, the distribution of the VNFs forming SFCs can impact the number of MECs needed, and consequently, the cost of provisioning 5G networks. Thus, the work in the thesis investigated the impact of the cost of different distribution of VNFs forming SFC requests for the MEC placement problem: i) centralized placement where SFC requests are fully hosted in a single slice of an MEC, and ii) distributed SFC assumption in which the VNFs forming an SFC are hosted in different number of MECs.

The contributions of the thesis are the following:

1. *A new enhanced share protection scheme, named* $1:N:K$.
2. *Evaluation of the impact of VNF distribution on the cost of provisioning a reliable 5G network with edge computing nodes.*
3. *Evaluation of the impact of traditional dedicated, shared and the new proposed enhanced shared protection schemes on the cost of provisioning a reliable 5G network with edge computing nodes;*

The relevance of the contributions of the thesis is related to the deployment of MEC nodes that should be carefully planned so that the low-latency and high-reliability requirements can be met simultaneously, since the replacement of a failed MEC by another one in a different location may lead to a violation of latency requirements. Most of the work on the support of 5G requirements neglects the reliability constraints and assumes that latency requirements can be supported regardless of any failure that might happen, which is an unrealistic assumption.

The MEC placement problem with a protection scheme is a constrained bi-objective optimization NP-hard problem. The problem of locating a reliable MEC was formulated as a generalization of the classical facility location problem, [Chantre and da Fonseca 2018]. All the feasible instances of this formulation were solved exactly using the Gurobi Optimizer solver. The metaheuristic NSGA-II was employed to solve the proposed formulation for networks larger than those with a feasible exact solution.

The remainder of this paper is organized as follows. Section 2 reviews related *work. Section 3 introduces the MEC placement problem with the assumption that the VNFs forming SFC requests are located in a single slice that is deployed on one MEC. Section 4 presents the contributions of the thesis considering the MEC placement problem with SFCs hosted on distributed MECs assuming that the VNFs forming the SFC requests are distributed on different MEC nodes. Section 5 summarizes the main results of the thesis. Section 6 lists the publications of the thesis. Finally, Section 7 draws the conclusions.*

2. Related Work

The MEC placement problem, investigated in the thesis, considered where to place MECs given a distribution of VNFs in a SFC. The placement of virtual network functions can be seen as a generalization of the virtual network embedding problem, an *NP*-hard problem. The resources of MECs can be divided into cloud-slices [Medeiros et al. 2023], which are referred to interchangeably as 'cloud-slices' or 'slices' in the thesis. Some authors approach the network function placement problem by focusing only on QoS metrics while others also consider reliability. For instance [Taleb et al. 2015] proposed a VNF placement algorithm for virtual 5G networks to minimize the length of the paths between users and their associated data anchor gateways as well as to optimize their session's mobility. The work in [Laghrissi and Taleb 2019] investigated VNF placement strategies, emphasizing potential issues that may disrupt VNF placement. The paper classifies the existing placement solutions based on the type (online or offline) and reliability-awareness and discusses metrics and objectives.

Redundancy is an essential aspect of service provisioning so that service resiliency can be assured. Service can be impacted by a failure of VNFs running on a VM, or by the failure of the physical machine hosting a VM. In [Chantre and da Fonseca 2017], a redundancy series-parallel model was introduced to improve the reliability of LTE Evolved Multimedia Broadcast Multicast Services (eMBMS) services. In [Kherraf et al. 2019], the authors have considered a workload assignment problem based on latency and reliability requirements to decide to which MEC node a workload should be assigned. However, their method does not consider that MECs and slices to be selected can be prone to failure. The authors in [Yao and Ansari 2019] investigated the trade-off between reliability and cost for resource provisioning in fog-aided IoT networks. However, the work in [Yao and Ansari 2019] does not explicitly explain the means of preventing VM failures, nor does it consider the existing traditional protection schemes to provide redundancy, nor the levels of redundancy taken into consideration on the fog resource provisioning problem.

The work in the thesis differs from those in the literature since it formulated the MEC node placement problem considering the stringent reliability requirements and latency simultaneously. Besides, the work in the thesis also investigated the impact of the dedicated and shared protection scheme in providing a reliable MEC node placement problem. The thesis investigated the impact of the cost of providing reliability for different SFC distributions for the MEC placement problem. To the best of our knowledge, none of the existing work investigates the impact of the distribution of the VNFs forming the SFC requests on the cost of providing reliability in 5G networks. Therefore, the presented thesis expands the state-of-the-art in different manners.

3. MEC placement problem in protected 5G networks with SFCs hosted in one MEC

This part of the thesis considered a scenario of a centralized placement where a single MEC node fully hosts the VNFs forming an SFC request. The impact of hosting VNFs composing an SFC in a single MEC and the cost of providing a reliable 5G network is investigated. The provisioning of reliability in an NFV-based MEC deployment is critical since a failure of a slice or a MEC can cause a service outage, breaking the continuity of the hosted SFC. To protect the SFCs against failure and ensure the stringent reliability required by the 5G use cases, different protection schemes are investigated. The effectiveness of the traditional dedicated protection, the shared protection schemes, and the new enhanced shared protection scheme proposed by the authors is evaluated to provide reliability in 5G networks with SFC hosted in one MEC. We aim to evaluate the cost of provisioning 5G services requirements that the centralized SFC incurs in 5G networks in terms of the number of MECs deployed. The impact that each protection scheme employed in the centralized SFC model is also of paramount importance do devising the number of MECs and slices demanded for furnishing the low latency and reliability requirements of 5G services.

We formulated the optimization problem as a bi-objective non-linear. To solve the bi-objective formulation, we first minimize the number of slices and then the number of deployed MECs. The goal of minimizing the number of slices is to assign as many demand points as possible to a single slice. To achieve this, these slices need to be packed into the minimum number of active MEC nodes, while avoiding spreading them across multiple MECs. To solve this problem, bi-objective non-linear constrained mathematical formulations were proposed and exactly solved, and solutions were also derived by employing the non-dominated sorting genetic algorithm for large networks.

4. MEC placement problem in protected 5G networks with SFCs hosted on distributed MECs

This problem addressed in the thesis aims at evaluating the impact of VNF distribution on the cost of providing reliability in the design of 5G networks. The thesis first evaluated the impact of VNF distribution on the cost of providing reliability in the design of 5G networks by placing a VNF (i.e., the last VNF) of an SFC in a separate MEC. This approach was referred to as the partial placement of VNFs forming a SFC. Then the impact of the cost of VNF distribution by placing all VNFs forming an SFC in separate MECs referred to as a fully distributed placement was evaluated.

The effectiveness of the traditional dedicated protection, the shared protection schemes, and the newly proposed enhanced shared protection scheme to provide reliability requirements of 5G networks with SFC hosted on distributed MECs were evaluated. In the evaluation of the impact of VNF distribution hosted on different MECs is paramount to consider the types of SFC requests, and the types of VNFs. For each protection scheme, we evaluated the number of distributed MECs that lead to a cost higher when the VNFs of an SFC are hosted in one single MEC.

The goal is to find the optimal locations for a minimal number of MECs, i.e., to find the minimum number of MECs, that can serve all the demand points and position them appropriately to achieve the goal. Moreover, the number of slices must be minimized. The objective was to assess the cost impact of distributing virtual network functions (VNFs) on an InP while provisioning a reliable 5G network that is enabled with Mobile Edge Computing (MEC) nodes. The aim is to minimize the number of slices and MECs required for a 5G network design with minimum cost while meeting the stringent requirements of latency and high-reliability level. The MEC placement problem is formulated as a bi-objective nonlinear mathematical problem. A linearization of the non-linear formulation of the MEC placement problem is proposed to evaluate the precision of the proposed meta-heuristic.

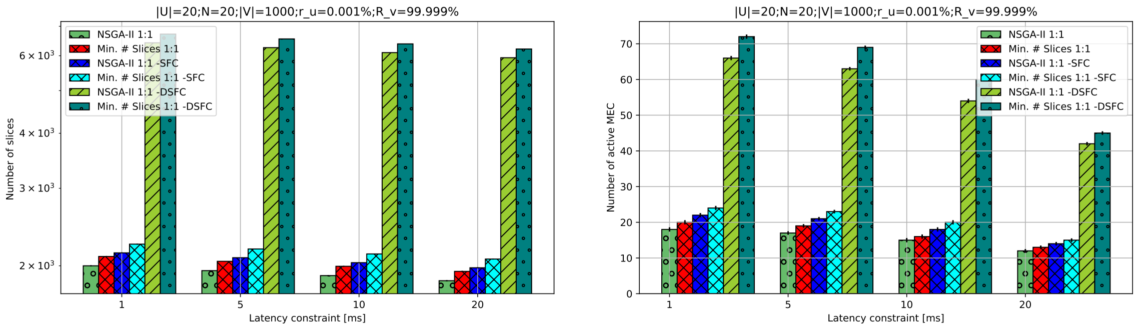
5. Results

In this section, the main results obtained in the thesis are presented. Detailed analyses can be seen in the thesis [Chantre and da Fonseca 2023]. The scenario for the MEC placement problem is composed of MEC nodes, distributed in a grid topology over an area of 1000 x 1000 meters. The non-dominated sorting genetic algorithm tailored to solve the MEC Location Problem was solved on a JMetalPy version 5.3. The linearized integer formulation of the MEC placement problem was solved with implemented in the Gurobi Optimizer solver.

For each of the protection schemes, dedicated $(1:1)$, shared $(1:N)$ and enhanced shared $(1:N:K)$ a comparison of the cost for the proposed SFC distribution in the MECs placement problem were evaluated. The non-linear bi-objective formulation with SFCs fully deployed in one MEC is labeled in the figures as $(NSGA-II)$, the partial distribution of SFC is labeled as $(NSGA-II\ SFC)$, and the fully distribution of SFC is labeled then as $(NSGA-II\ DSFC)$. The linearized bi-objective formulation of MEC placement problem is then labeled as follow (LP) for SFCs fully deployed in one MEC, $(LP\ SFC)$ for SFC partial distributed and $(LP\ DSFC)$ for the formulation that the SFCs are fully distributed across different MECs. A mono-objective formulation with $K=3$ level of redundancy is labeled as follow $(Min.\#Slice\ K=3)$ for the SFC fully deployed in one MEC, $(Min.\#Slice\ K=3\ SFC)$ for SFC partially distributed, and $(Min.\#Slice\ K=3\ DFC)$ for SFC fully distributed.

In Figures 1a and 1b, the number of demanded slices and MECs are evaluated as function of latency constraints respectively. From the definition of the $1:1$ protection schemes, every demand point has a dedicated secondary backup node reserved in case of failure. Thus, in a full distributed VNF location where each single VNF forming an SFC request is placed on separated MEC, the dedicated protection scheme employed triple the number of demanded slices and MECs respectively. The impact of distributing the VNFs in MEC placement problem with a dedicated protection scheme is very high in the order of 68 MECs compared to the 24 MECs in the partial distribution and the 18 MECs in the SFC fully deployed in a single MEC.

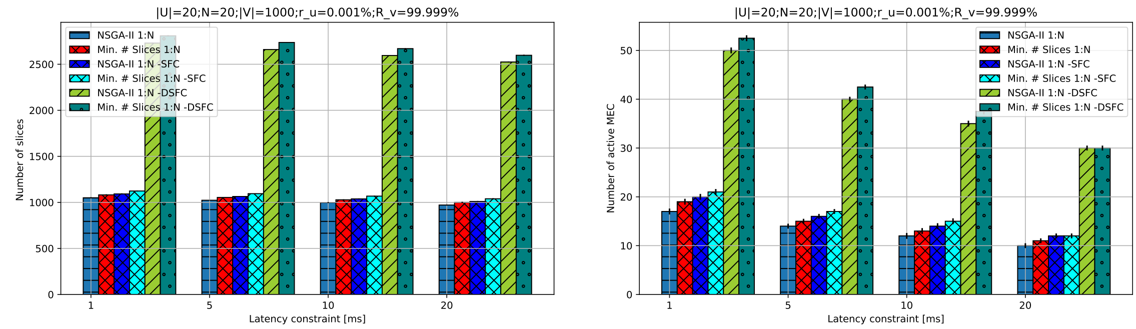
Figures 2a and 2b show the number of slices and MECs demanded as function of latency constraints. Figures 2a and 2b show that the variation in terms of latency constraint minimize the impact of the number of shared demand point N. The figures also confirm the observation that the full distributed approach may not be preferable to an InP, as the configuration of the 5G MEC based network in the fully distributed (NSGA-II 1:N DSFC) VNF placement demand more MECs, e.g in the order of 50 MECs compared to the 20 MECs in partial distributed VNFs (NSGA-II 1:N SFC).



(a) Number of Slices demanded as a function of latency constraints.

(b) Number of MECs demanded as a function of latency constraints.

Figure 1. Number of Slices and MEC demanded for a full distributed SFC with $1:1$.



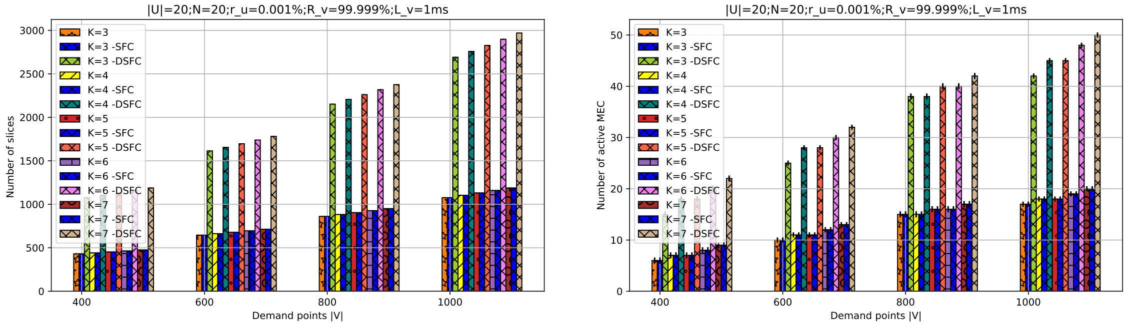
(a) Number of Slices demanded as a function of latency constraints.

(b) Number of MECs demanded as a function of latency constraints.

Figure 2. Number of Slices and MEC demanded for a full distributed SFC with $1:N$.

Figures 3a and 3b show the costs of demanded slices and MECs as a function of demand points $|V|$ from $K = 3$ to $K = 7$ for the VNFs of a SFC placement studied in this thesis work. The results show that the costs increases as the number of demand points increases. We also observe that as we increase the number of K redundant nodes so the costs in terms of demanded slices and MECs increases. So the full distribution of VNFs of a SFC placement in which each VNF belonging a SFC request must be placed on a separated VM results in a significant impact in terms of costs and operation of the network for an InP. Therefore, comparing the full distribution of VNFs placement with the VNFs forming a SFC request placed in one VM and the partial distributed (e.g. the last VNF of an SFC placed on a separated VM), the later approach might be preferable as they provide twice less slices and MECs when comparing with the former.

Results demonstrated that when end-users demand a higher level of redundancy, more resources are needed, thus increasing the cost linearly with the level of redundancy. For the centralized placement, the configuration with $1 : N : 3$ with two redundant nodes is cost-beneficial for end-users as it provides an extra level of protection with the same cost of the shared $1 : N$ protection scheme. The type of protection used in the placement of Multi-access Edge Computing (MEC) nodes depends on the criticality of



(a) Number of Slices required. **(b) Number of MECs required.**

Figure 3. Number of slices and MECs required as a function of demand points for a full distributed SFC with $1:N:K$.

the service being provided. For instance, services such as ultra-Reliable Low Latency Communications (uRLLC) that are used for applications such as remote surgery may need dedicated nodes to ensure service continuity in case of failures. In this case, a $1:1$ protection scheme may be preferable, despite the high costs. On the other hand, services like autonomous vehicles can benefit from an additional level of redundancy in case of node failure along a route, and the $1:N:K$ protection scheme may be a better option, as it provides an extra level of redundancy at a lower cost compared to other protection schemes.

6. Publications

Results of the thesis were published as papers in 3 journal papers, 1 magazine paper and 1 conference paper: [Chantre and Fonseca 2022] in the IEEE Transactions on Network and Service Management (Impact Factor 5.3), [Chantre and Saldanha da Fonseca 2020] in the IEEE Journal on Selected Areas in Communications (Impact Factor 16.4), [Chantre and da Fonseca 2018] in the IEEE Journal on Selected Areas in Communications (Impact Factor 16.4), [Chantre and d. Fonseca 2018] in the IEEE Communications Magazine (Impact Factor 11.2), [Chantre and da Fonseca 2017] in the IEEE International Conference on Communications.

7. Conclusions

The thesis reviewed in this paper explored different protection schemes for furnishing reliable services in 5G MEC-based networks. Moreover, we evaluated the impact of the cost of different distribution methods for VNFs of SFC requests in the MEC placement problem with protection schemes. The findings of this thesis suggested that fully distributing VNFs results in a more expensive network design compared to the other methods, regardless of the protection scheme used. Therefore, depending on the requirements of the end users one may prefer the partial or centralized distribution methods with an enhanced protection scheme as they require fewer nodes and slices while providing an additional level of redundancy compared to other protection schemes. This thesis differs from the existing works in the literature given that the majority assumes that the VNFs are distributed and the costs and the impacts of the VNF distribution are neglected. This thesis studied the impact of VNF distribution on the cost of providing reliability in 5G network design. The thesis evaluated the problem of locating MECs and slices in a 5G infrastructure are protected by different protection schemes and yet the thesis proposed a new enhanced scheme $1 : N : K$ to protect 5G networks. Possible future directions of this work can be explored for disaster recovery, and the dynamic VNF placement approaches including multiple self-contained networks over a shared infrastructure, enabled by self-optimization and self-healing functionalities.

Acknowledgements

This study was financed in part by PEC-PG-CAPES/CNPq.

References

Chantre, H. and d. Fonseca, N. L. S. (2018). Reliable Broadcasting in 5G NFV-Based Networks. IEEE Communications Magazine, 56(3):218–224.

Chantre, H. D. and da Fonseca, N. L. S. (2017). Redundant placement of virtualized network functions for LTE evolved Multimedia Broadcast Multicast Services. In 2017 IEEE International Conference on Communications (ICC), pages 1–7.

Chantre, H. D. and da Fonseca, N. L. S. (2018). Multi-Objective Optimization for Edge Device Placement and Reliable Broadcasting in 5G NFV-Based Small Cell Networks. IEEE Journal on Selected Areas in Communications, 36(10):2304–2317.

Chantre, H. D. and da Fonseca, N. L. S. (2023). MEC Placement Problem in Protected 5G Networks. PhD thesis, Universidade Estadual de Campinas.

Chantre, H. D. and Fonseca, N. L. S. d. (2022). Design of 5g mec-based networks with 1:n:k protection scheme. IEEE Transactions on Network and Service Management, 19(3):2478–2491.

Chantre, H. D. and Saldanha da Fonseca, N. L. (2020). The Location Problem for the Provisioning of Protected Slices in NFV-Based MEC Infrastructure. IEEE Journal on Selected Areas in Communications, 38(7):1505–1514.

Kherraf, N., Sharafeddine, S., Assi, C. M., and Ghrayeb, A. (2019). Latency and Reliability-Aware Workload Assignment in IoT Networks With Mobile Edge Clouds. IEEE Transactions on Network and Service Management, 16(4):1435–1449.

Laghrissi, A. and Taleb, T. (2019). A Survey on the Placement of Virtual Resources and Virtual Network Functions. IEEE Communications Surveys Tutorials, 21(2):1409–1434.

Medeiros, A., Maio, A. D., Braun, T., and Neto, A. (2023). Tenet: Adaptive service chain orchestrator for mec-enabled low-latency 6dof virtual reality. IEEE Transactions on Network and Service Management, pages 1–1.

Taleb, T., Bagaa, M., and Ksentini, A. (2015). User Mobility-Aware Virtual Network Function Placement for Virtual 5G Network Infrastructure. In 2015 IEEE International Conference on Communications (ICC), pages 3879–3884.

Yao, J. and Ansari, N. (2019). Fog Resource Provisioning in Reliability-Aware IoT Networks. IEEE Internet of Things Journal, 6(5):8262–8269.