Designing and Evaluating a high-reliable and security-aware Identity and Access Management for Slicing Architectures

Rodrigo Moreira[1], **Tereza C. M. B. Carvalho**[2], **Flávio de Oliveira Silva**[3]

[1]Federal University of Viçosa (UFV)
38.810-000 – Rio Parnanaíba – MG – Brazil

[2]University of São Paulo (USP)
05.508-010 – São Paulo – SP – Brazil

[3]Federal University of Uberlândia (UFU)
38.408-100 – Uberlândia – MG – Brazil

rodrigo@ufv.br, terezacarvalho@usp.br, flavio@ufu.br

Abstract. *Network slicing architectures are fundamental for providing connectivity to demanding users and applications in heterogeneous network infrastructures. Such architectures have evolved significantly in recent years, especially with improvements in security and reliability functions. However, the improvements in these architectures are functionally specific and are not considered throughout the entire architecture lifecycle, opening opportunities for secure, and reliable native architectures. Thus, this paper designs and evaluates an Identity and Access Management (IAM) mechanism while providing security and reliability for building blocks of slicing architectures. Our findings concern a comparative evaluation of the IAM mechanism and its behavior under stress loads, as well as an experimental assessment of a secure defense mechanism against Distributed Denial-of-Service (DDoS) attacks.*

1. Introduction

Network slicing has revolutionized connectivity for users and applications on modern networks. Different state-of-the-art approaches made efforts to support multi-tenant network sharing and open up business opportunities across various verticals. These initiatives focused on building and evaluating architectures capable of deploying network slices in different domains, especially using heterogeneous technologies [Moreira et al. 2021]. Artificial Intelligence (AI) has played an important role in improving various aspects of the network slice lifecycle with security and reliability.

However, these improvements are separate from the architectural building blocks of network slicing and are incorporated into specific modules or functions [Debbabi et al. 2022]. The 3rd Generation Partnership Project (3GPP) specifies a security architecture for 5G Systems that includes authentication and key agreement to ensure secure message exchanging between slicing nodes [3GPP 2020]. There is a need for mechanisms to deal with security throughout the slicing lifecycle. In this paper, we designed and evaluated a general-purpose Identity and Access Management (IAM) mechanism for future network slicing architectures.

Among the main contributions of this paper, the following stand out: (1) an AuthToken issuing mechanism for network slicing blocks; (2) an empirical evaluation of IAM mechanisms deployed on a high availability Kubernetes environment; (3) an evaluation of the traffic mirroring mechanism in high availability environments for security network traffic monitoring.

The remainder of this paper is organized as follows: Section 2 presents existing works, Section 3 presents the contribution rationale, while Section 4 presents technical and evaluative details of our mechanism. Section 5 presents insights, and we draw conclusions from our experimental evaluation in Section 6.

2. Existing Work

This section presents and discusses relevant works that propose security enhancements for network slicing architectures. Some works implement security as a specific feature of slicing and not as an architectural entity that provides secure defense mechanisms. Others propose security for applications that run on network slicing, others propose security mechanisms between the functional blocks of network slicing architectures.

[Porambage et al. 2019] proposed a security keying scheme based on a combination of key generation, distribution, and management techniques for a Smart Factory use case. The scheme uses Shamir's secret sharing method to enable data access for third-party monitoring tools of network slicing. With this, it is ensured that third-party tools only monitor the correct network slice. Unlike this approach, our proposal for secure improvements encompasses mutual authentication of the entities that make up the slicing architecture through IAM.

In [Kiyemba Edris et al. 2020] proposed and evaluated Federated Identity Management (FIdM), which allows the sharing of identity information across multiple domains and service providers while preserving privacy and security through mutual authentication, authorization, identity protection, secure access, and interoperability between domains. They present use cases for FIdM in 5G networks such as enabling secure and seamless roaming across different networks and service providers. In our method, we propose a distributed identity mechanism where functional blocks of the slicing architecture are equipped with a Distributed Denial-of-Service (DDoS) defense mechanism.

[Wijethilaka and Liyanage 2022] proposed an L5GO framework for security-oriented network slicing, where resource management mechanisms exist while providing security features for network slices. The proposed framework provides security to be considered in the network slicing process; each network slice can have a security enhancement that resides in the Security Function Repository.Alternatively, we propose a closed-loop approach for security enhancement in each building block of a network slicing architecture.

3. Proposed Architecture

We propose an IAM mechanism for managing identities and access in network slicing architectures that provides defense features against denial-of-service attacks using distributed Machine Learning (ML). Our solution uses ML-Agents to notify the IAM when a functional block is under attack [Moreira et al. 2023]. A denial-of-service attack on slicing architectures aims to compromise service availability by flooding it with fake requests.

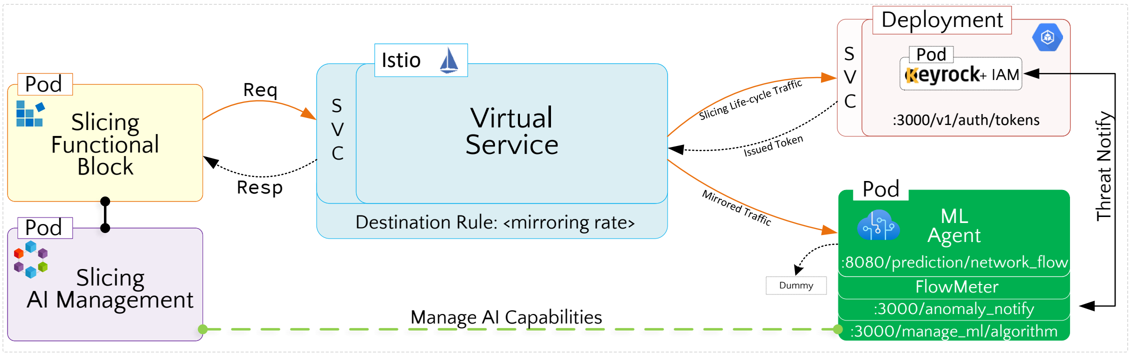


Figura 1. Proposed Architecture Overview.

In slicing orchestration architectures on Kubernetes-based domains such as Future Internet Brazilian Environment for Experimentation (FIBRE) (New Generation), all operations and communication between functional blocks should require the exchange of session tokens. We have extended Keyrock to add IAM functionality as a security center for future slicing architectures. Functional blocks exchange messages using the session tokens requested to a Virtual Service built with Istio technology, which enables traffic mirroring in Kubernetes.

When the Virtual Service receives a token-issuing request, it forwards it to the IAM, which authenticates the functional block and issues an Auth-Token for communication with other blocks. The Virtual Service forwards all mirrored traffic to the ML-Agent through a Destination Rule while IAM handles the token request.

The ML-Agent is a closed-loop for cognitive and security services that proactively implements security for functional blocks of slicing architectures. It uses FlowMeter [Lashkari et al. 2017] technology to sample flows and submit them to an API for predicting network traffic anomalies. If a flow matches a DDoS attack pattern, the ML-Agent notifies the IAM so that the architecture administrator can make a decision. The ML-Agent can be equipped with different skills and training to offer cognitive and security services.

The ML models in the ML-Agent can be updated at any time by the Slicing AI Management block through an Application Programming Interface (API). The architecture admin enables the ML-Agent with up-to-date models. The actual responses of the mirrored traffic that the ML-Agent targets are discarded, with only the passive analysis of the network traffic destined for the IAM taking place, implying no overhead for the service.

4. Experimental Setup

To validate the applicability of the proposed approach, we devised an experimental scenario that proposes to functionally assess the capacity and behavior of the IAM against different workloads[1]. In our approach, we have the assumption that all functional blocks of slicing are equipped with the ML-Agents that receive all traffic through traffic mirroring. Thus, the ML-Agent can notify the IAM Security Center if a denial-of-service

[1] Available at `https://github.com/romoreira/Keyrock-IAM.git`

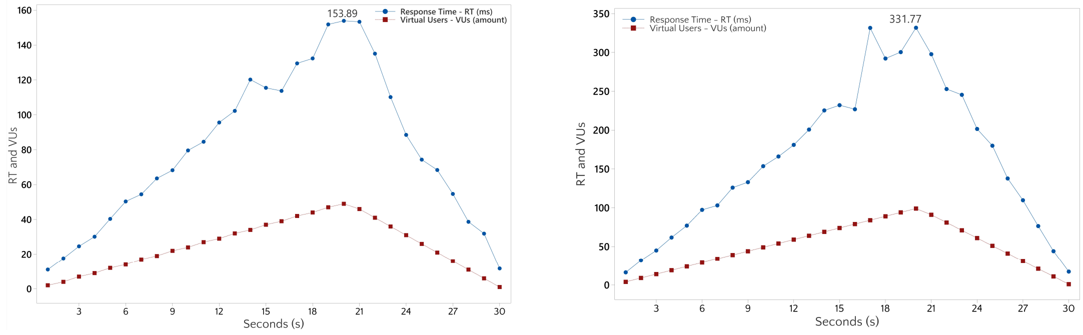
attack occurs. For this experimentation, we scripted benchmarking scenarios using the K6 benchmark tool deployed in a Kubernetes environment.

We built an experimental scenario on a two-node Kubernetes cluster in Azure to test the IAM using a MySQL database and ML-Agents equipped with trained ML models. We set a 99% reliability metric and measured Response Time (RT) while increasing the number of Virtual Users (VUs) directing requests to the IAM. Futhermore, our scenario consists of rising the number of VUs that direct requests to IAM increasing the number of VUs in each experiment according to 50, 100, 1000, 5000, and 10000.

ML-Agent was requested to predict network flows regarding benign classes or DDoS attacks, namely: DrDoS-DNS, DrDoS_MSSQL, DrDoS_NetBIOS, DrDoSa_SNMP, DrDoS_UDP, Syn, TFTP, UDP-lag. These names refer to the classes of DDoS attacks as it is on dataset provided by [Sharafaldin et al. 2019]. Thus, we previously trained the K-Nearest Neighbors (KNN), Multilayer Perceptron (MLP), Random Forest (RF), and Support Vector Machine (SVM) algorithms and exported the model as a *pickle* file for later use by the API of ML-Agent. The accuracy of these models was over 90% and in this paper, they are exclusively intended to empower ML-Agent with on-the-fly prediction.

5. Evaluation and Discussion

We performed experiments to evaluate the ability of IAM to issue Auth-Tokens and respond to requests while under attack. We used one (1) replica to measure the maximum capacity of the slicing functional block before communication errors occurred.



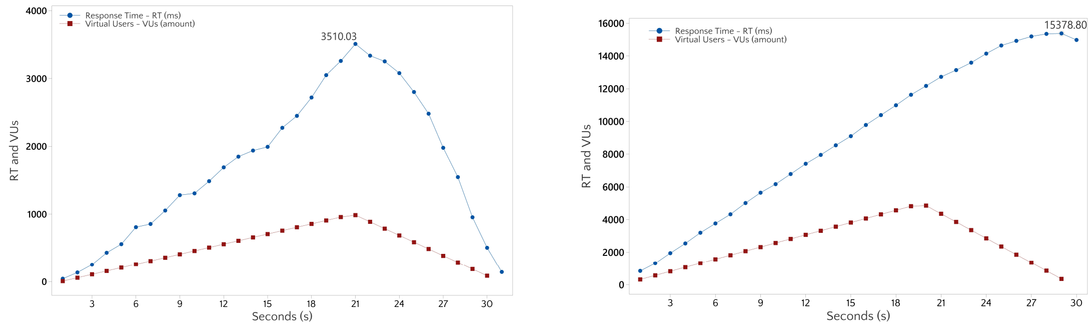
(a) 50 VUs.

(b) 100 VUs

Figura 2. Increasing VUs and requests towards IAM for 50 and 100 VUs.

We increased the number of users, who requested the Auth-token to the IAM. For 50 and 100 VUs, according to Figure 2, it can be seen that until the initial 20 seconds the number of VUs increased to the one configured for the experimental round, and it can be seen that the RT increases proportionally to the number of VUs.

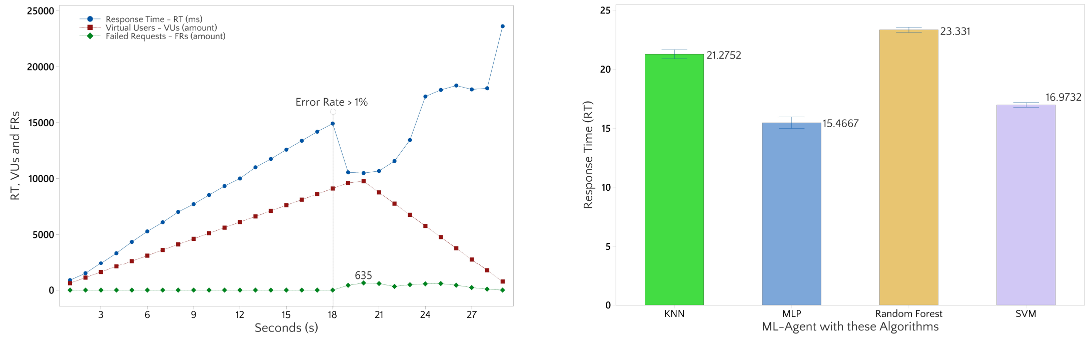
Note that for 1000 VUs and 5000 VUs, as shown in Figure 3, no errors are recorded in the request due to the high workload. However, for 10000 VUs after 18 seconds, with above 9117 VUs, reliability violations start as represented by Figure 4a. This leads us to admit that an IAM service for slice orchestration with security mechanisms should be deployed in a scalable and high-availability environment like Kubernetes, which allows configuring replication factors based on workload.



(a) 1000 VUs

(b) 5000 VUs

Figura 3. Increasing VUs and requests towards IAMfor 1000 and 5000 VUs.



(a) 10000 VUs

(b) ML-Agent RT for one (1) Virtual User (VU).

Figura 4. Increasing VUs and requests towards IAM and breakdown reliability.

We measured the ML-Agent to respond to network flow prediction requests regarding its traffic class. The K6 tool submitted a network flow instance for 30 seconds where we could verify the RT of the ML-Agent for some machine learning algorithms. All these machine learning algorithms were requested to respond regarding the traffic class of a given network flow. Thus, according to Figure 4b it can be seen that when the ML-Agent was equipped with the trained ML algorithms, it demanded the smallest RT (about $15.46ms$) to predict network flow for benign or DDoS attacks considering a 95% confidence interval, while RF took the longest time (about $23.33ms$) to predict a network flow for its traffic class.

When implementing slicing architectures with security features, it is important to consider the RT of each algorithm to ensure optimal performance and reliability. Using high-availability platforms such as Kubernetes to host slicing architecture blocks may be appropriate.

6. Concluding Remarks

In this paper, we have designed and evaluated an IAM for future slicing architectures that meet security and high-reliability requirements. Predominantly, these requirements are implemented as specific functionalities in functional blocks rather than fully driving the building block of slicing architectures. Through our evaluation in high-demand scenarios, we demonstrated that our IAM is capable of handling a high workload, with a single Pod delivering reliable performance. Finally, we found that our ML-Agent responds readily to network flow prediction workloads, highlighting the potential of AI in enhancing the security and reliability of network slicing architectures.

In conclusion, we validated the traffic mirroring mechanism to monitor the slicing architectures for potential DDoS attacks. This method offers a promising approach to enhance security, and future investigations should explore its performance when deployed with the horizontal pod auto-scaler enabled and checking system limits. Additionally, further research is needed to explore other intelligent defense mechanisms against security attacks.

Acknowledments

The authors thank the Research Support Foundation of the State of São Paulo (FAPESP) under grant number 2018/23097-3 for the thematic project Slicing Future Internet Infrastructures (SFI2).

Referências

3GPP (2020). Security architecture; procedures for 5G System. Technical specification (ts). TS 33.501 V16.2.0 (2020-03).

Debbabi, F., Jmal, R., Chaari, L., Aguiar, R. L., Gnichi, R., and Taleb, S. (2022). Overview of AI-based Algorithms for Network Slicing Resource Management in B5G and 6G. In *2022 International Wireless Communications and Mobile Computing (IWCMC)*, pages 330–335.

Kiyemba Edris, E. K., Aiash, M., and Loo, J. K.-K. (2020). The case for federated identity management in 5G communications. In *2020 Fifth International Conference on Fog and Mobile Edge Computing (FMEC)*, pages 120–127.

Lashkari, A. H., Draper-Gil, G., Mamun, M. S. I., Ghorbani, A. A., et al. (2017). Characterization of tor traffic using time based features. In *ICISSp*, pages 253–262.

Moreira, R., Martins, J. S. B., Carvalho, T. C. M. B., and de Oliveira Silva, F. (2023). On enhancing network slicing Life-Cycle through an AI-native orchestration architecture. In *The 37-th International Conference on Advanced Information Networking and Applications (AINA-2023) (AINA-2023)*, Federal University of Juiz de Fora, Brazil.

Moreira, R., Rosa, P. F., Aguiar, R. L. A., and de Oliveira Silva, F. (2021). NASOR: A network slicing approach for multiple Autonomous Systems. *Computer Communications*, 179:131–144.

Porambage, P., Miche, Y., Kalliola, A., Liyanage, M., and Ylianttila, M. (2019). Secure keying scheme for network slicing in 5G architecture. In *2019 IEEE Conference on Standards for Communications and Networking (CSCN)*, pages 1–6.

Sharafaldin, I., Lashkari, A. H., Hakak, S., and Ghorbani, A. A. (2019). Developing realistic distributed denial of service (ddos) attack dataset and taxonomy. In *2019 International Carnahan Conference on Security Technology (ICCST)*, pages 1–8.

Wijethilaka, S. and Liyanage, M. (2022). A novel network slicing based security-as-a-service (SECaaS) framework for private 5G networks. In *2022 IEEE Latin-American Conference on Communications (LATINCOM)*, pages 1–6.