5G-WAVE: A Core Network Framework with Decentralized Authorization for Network Slices

Pragya Sharma*, Tolga Atalay*, Hans Andrew Gibbs†, Dragoslav Stojadinovic†,
Angelos Stavrou*†, Haining Wang*
*Bradley Department of Electrical and Computer Engineering, Virginia Tech, USA
†Kryptowire LLC, McLean, VA, USA
Email: *{pragyasharma, tolgaoa, angelos, hnw}@vt.edu, †{hgibbs, dstojadinovic, angelos}@kryptowire.com

IEEE INFOCOM 2024 - IEEE Conference on Computer Communications | 979-8-3503-8350-8/24/$31.00 ©2024 IEEE | DOI: 10.1109/INFOCOM52122.2024.10621131

***Abstract*—5G mobile networks leverage Network Function Virtualization (NFV) to offer services in the form of network slices. Each network slice is a logically isolated fragment constructed by service chaining a set of Virtual Network Functions (VNFs). The Network Repository Function (NRF) acts as a central OpenAuthorization (OAuth) 2.0 server to secure inter-VNF communications resulting in a single point of failure. Thus, we propose 5G-WAVE, a decentralized authorization framework for the 5G core by leveraging the WAVE framework and integrating it into the OpenAirInterface (OAI) 5G core. Our design relies on Side-Car Proxies (SCPs) deployed alongside individual VNFs, allowing point-to-point authorization. Each SCP acts as a WAVE engine to create entities and attestations and verify incoming service requests. We measure the authorization latency overhead for VNF registration, 5G Authentication and Key Agreement (AKA), and data session setup and observe that WAVE verification introduces 155ms overhead to HTTP transactions for decentralizing authorization. Additionally, we evaluate the scalability of 5G-WAVE by instantiating more network slices to observe 1.4x increase in latency with 10x growth in network size. We also discuss how 5G-WAVE can significantly reduce the 5G attack surface without using OAuth 2.0 while addressing several key issues of 5G standardization.**

***Index Terms*—5G, Network Slicing, Decentralized Authorization, WAVE, Microservices**

I. Introduction

The deployment of Fifth-Generation (5G) mobile networks is an ongoing process that is gaining momentum with each passing year. As part of their service offerings, 5G mobile networks have to meet the strict Quality of Service (QoS) requirements for a diverse set of use cases. To achieve such versatility, 5G services are offered as logically isolated end-to-end fragments known as network slices. This approach enables 5G operators to create independent network instances that can be tailored for different sets of requirements.

Compared to Long Term Evolution (LTE), delivery of services through network slicing relies on Network Function Virtualization (NFV) for increased flexibility. Leveraging NFV, network slices are constructed by service chaining a set of Virtual Network Functions (VNFs) hosted on Commercial Off-The-Shelf (COTS) hardware. This allows operators to scale their deployments more efficiently and reduce costs by replacing dedicated service hardware. The network slice VNFs communicate with one another through HTTP transactions within a Service-Based Architecture (SBA) [1]. A key feature of the 5G core SBA is the use of standardized Application Programming Interfaces (APIs) to enable interoperability between alternative implementations of the 5G core VNFs. The individual HTTP transactions between the VNFs are secured using an OpenAuthorization (OAuth) 2.0 server that distributes access tokens for moderating access requests [2].

According to the 3rd Generation Partnership Project (3GPP) [2], [3] standardizations, the Network Repository Function (NRF) maintains the metadata profiles of all the VNFs and supports mutual discovery operations that establish connections between service providers and requesters in a given administrative domain. 3GPP also defines NRF as an OAuth 2.0 server that distributes access tokens to enable inter-VNF authorization for the secure consumption of services.

Motivation. While this fundamental approach to service discovery and authorization fulfills the basic requirements for security, it promotes a rigid framework for network slice construction where VNFs rely on a central trusted entity. In a scenario where the NRF is compromised, an adversary can issue malicious tokens or tamper service contracts for the provider VNFs. Moreover, the 5G SBA is a distributed ecosystem deployed as microservices, instead of monolithic VNFs. Thus, the dependency on a central entity for authorization can lead to performance bottlenecks in control plane signaling.

To address this security gap, we propose an integrated platform called **5G-WAVE** for network slice construction with decentralized authorization among the 5G core VNFs. WAVE [4] is a decentralized authorization and verification framework which enables transitive delegation of access among entities without relying on a central trust authority. It allows authorization between *'WAVE entities'* by creating contracts called *'attestations'* for resource access. When service requests arrive, WAVE uses these attestations to verify their permissions.

To integrate WAVE seamlessly into the 5G ecosystem, we propose a microservice-based augmentation to a Kubernetes-based 5G core deployment. The WAVE entities are deployed adjacent to each VNF as Side-Car Proxies (SCP) [5]. Using the SCPs, an indirect communication service mesh [6] is created around the 5G core to offload the authorization functionality from the 5G core VNF to the WAVE infrastructure. The SCPs serve as WAVE entities granting attestations for resource access, thereby removing the reliance on a central token-

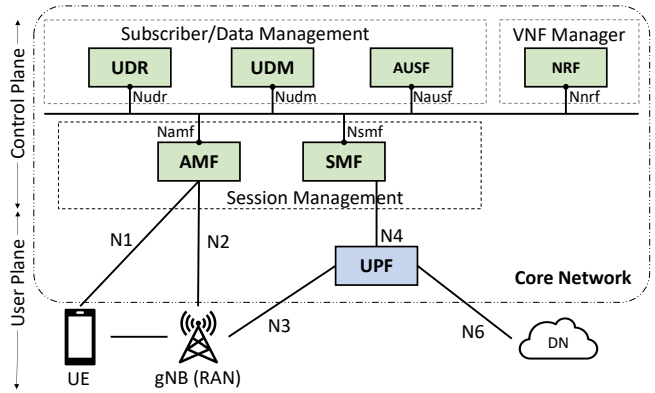

Fig. 1: 5G core service-based architecture

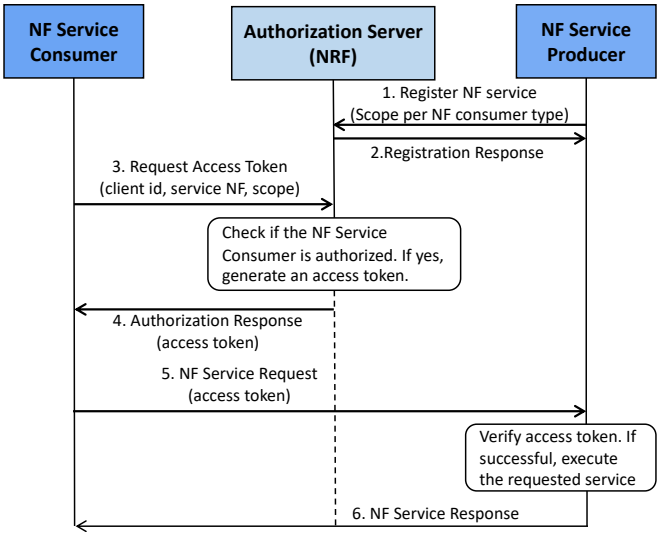

Fig. 2: NRF as OAuth2.0 server for inter-VNF authorization

issuing server for authorization. To facilitate implementation and integration, we use the WAVE SCP (wSCP) in conjunction with a redirection SCP (rSCP) to create an interception/verification pipeline for the 5G core service requests. If the service request is verified successfully, it is forwarded to its original destination, i.e., the target 5G core VNF. In our evaluations, we measure the latency overhead introduced to the 5G operations because of integrating WAVE into the core network. The results show that an authorization latency of 155ms is introduced to individual HTTP transactions. As we scale the slices, we observe a 1.4x increase in authorization latency with 10x network growth. On the other hand, the 5G-WAVE integration addresses several key issues from the 3GPP network slicing security enhancement efforts. Furthermore, by proposing an alternative to OAuth 2.0, multiple attack vectors that leverage the OAuth threat surface are eliminated.

The main contributions of this work are summarized below.

- We propose the 5G-WAVE integrated platform for decentralized inter-VNF authorization.
- We implement the authorization flow using the OpenAir-Interface (OAI) [7] 5G core and test it using gNBSIM [8]. An interception/verification pipeline is constructed using two SCPs to facilitate the seamless integration of WAVE into the 5G core.
- We measure the latency overhead of the proposed platform in comparison to native 5G core with single and multiple network slices.
- A security analysis is conducted to demonstrate how the proposed 5G-WAVE integration addresses key security issues from the 3GPP standardization as well as the OAuth attack surface.

II. Background

A. 5G Core Overview

The 5G core has a SBA comprised of a set of interconnected VNFs as shown in Figure 1. Each VNF needs to be properly authorized to access the resources of another [1]. The key network functions of the 5G core are summarized as follows: 1) Network Repository Function (NRF) serves as a centralized VNF profile repository and facilitates mutual discovery and authorization. 2) Unified Data Repository (UDR) serves as a centralized repository for subscriber data. 3) Unified Data Management (UDM) manages the subscriber context data and authentication credentials. 4) Authentication Server Function (AUSF) terminates the 5G Authentication and Key Agreement (AKA) service chain in the 5G home network. 5) Access and Mobility Management Function (AMF) manages the subscriber access and mobility within the network. 6) Session Management Function (SMF) manages the subscriber data session within the network. 7) User Plane Function (UPF) handles the user traffic by forwarding it between the user and Data Nodes (DNs).

B. 5G Core Authorization - Current Standard

3GPP has standardized OAuth 2.0 [2], [3] as the default authorization mechanism between VNFs. It denotes the NRF as an OAuth 2.0 server. Any other VNF Service Consumer (aka service requester) acts as an OAuth 2.0 client and the VNF Service Producer (aka service provider) acts as an OAuth 2.0 resource server. As illustrated in Figure 2, the service producer registers with the NRF as an OAuth 2.0 resource server along with the scope of authorization for service consumers. When the service consumer needs to access the services of the producer, it sends an authorization request to the NRF along with the requested scope and relevant access data of the target provider. The NRF checks whether the consumer is authorized to access the requested service. If it is authorized, then the NRF generates an access token with the appropriate scope and forwards it to the consumer as a response. After authorization, the consumer sends a service request to the provider with the access token. The service provider validates the received access token and responds with the requested resource.

C. Side-car Proxy and Service Mesh

The foundation of our 5G-WAVE integrated platform is built upon the concept of SCPs illustrated in Figure 3. In a Kubernetes cluster, the smallest logical unit of deployment is

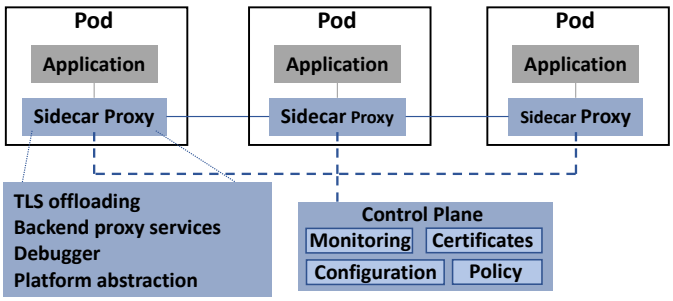

Fig. 3: Side-car proxy overview in Kubernetes environment

a pod. A single pod can be made up of multiple containers that share the same network sandbox.

The SCP is a non-functional container in a pod that provides the main application with a desired set of abstractions for non-functional services such as security procedures, configurations, policies, and monitoring. Combined with a service mesh that connects the SCPs, the deployments can be managed from a centralized control plane. Given their flexible nature, SCPs have become a popular way of augmenting applications with third-party functionality. Furthermore, in 3GPP Release 16, SCPs have been standardized as a method of indirect communication between the 5G core VNFs [1].

D. Authorization in WAVE

WAVE is an authorization framework designed for distributed systems to handle transitive delegations and revocations without having a central trust authority [4]. Initially designed for distributed systems like IoT systems in the built environment [9], [10], WAVE can be used as a general-purpose decentralized authorization framework. Based on existing decentralized Trust Management (TM) systems, WAVE improves them with additional security guarantees. The delegations and revocations of permissions are cryptographically enforced, offering confidentiality and invisibility to untrusted parties. If an intrusion occurs in the authorization server, the WAVE architecture ensures that the permissions of users who are not compromised remain intact, and the intruder is unable to view any existing system permissions. WAVE achieves these security guarantees by implementing 1) a Graph-Based Authorization (GBA) model, 2) Reverse Discoverable Encryption (RDE) of attestations, and 3) horizontally scalable untrusted storage.

The WAVE design relies on a GBA model for representing transitive delegations, where the state of the system is maintained in the form of a perspective graph for each entity of the network. Each WAVE *entity* is associated with a unique public-private key pair, representing a user in the system. WAVE achieves decentralized trust by allowing only the transacting entities to grant or revoke permissions through encrypted attestations, which are signed certificates from the *issuer/authorizer* entity to the *subject/requester* entity. Entities form the vertices and attestations form the edges between the authorizer and the requester in the WAVE authorization graph [11]. An entity which has been granted delegated permissions can form a WAVE proof by traversing the path in the authorization graph from itself to the authorizing entity, discovering the attestations associated with it. Furthermore, any user can verify this WAVE proof, unlike other decentralized trust management systems like SDSI/SPKI [12] and Macaroons [13], thus removing reliance on a central verification authority. WAVE objects are stored in horizontally scalable distributed servers to enforce integrity. WAVE has a novel Unequivocable Log Derived Map (ULDM) design for the storage, which scales better than blockchain for addition and retrieval of objects.

Within the scope of our study, we are not concerned with storage mechanisms of WAVE. Our evaluation is focused on the performance of WAVE as a decentralized authorization platform for supporting the 5G core SBA. All operations on the client side are handled by a 'WAVE daemon' which acts as an agent to create entities, attestations and proofs, and publishes/retrieves objects to/from persistent storage.

III. 5G-WAVE Integrated Platform

The goal of 5G-WAVE is to enable 5G core VNFs to leverage WAVE for mutual authorization via permission delegations. This section provides a detailed overview of the 5G-WAVE system. Following that, we describe inter-pod message exchange, intra-pod flow details (wSCP and rSCP roles), and implementation specifics.

A. 5G-WAVE Integrated System Overview

The overview of the 5G-WAVE integration is illustrated in Figure 4. Taking place in a Kubernetes cluster, each key component is deployed either at the pod- or cluster-level. To facilitate the interpretation, the components are grouped into three categories: 5G core VNFs, 5G-WAVE integration agents and finally the native WAVE elements.

As illustrated in Figure 4, a single pod is comprised of three different containers. These are the 5G core VNF containers and the 5G-WAVE integration agents (i.e., rSCP and wSCP). The 5G core VNFs are the native OAI software entities [14] that are the primary applications of interest in our deployment. To adapt them for this integration, we modified their instantiation process so that they trigger the internal flow of the wSCP (WAVE SCP) before becoming functional.

The rSCP (Redirection SCP) is a custom interception proxy that performs message redirection on incoming HTTP requests. It interacts with the wSCP during the WAVE authorization chain to seek validation on incoming API calls. While the rSCP is a helper to intercept the HTTP messages, wSCP provides functional security features to verify the authorization of the incoming service requests. The wSCP executes the functionalities of a WAVE client on behalf of its respective 5G core VNF. This includes the creation of the WAVE entity as well as the attestations.

The remaining two components are the WAVE daemon and the storage server, both of which are native to WAVE and are deployed at the cluster level. A WAVE daemon serves as an agent responsible for generating entities and attestations and interacts with wSCPs. A WAVE storage server functions as

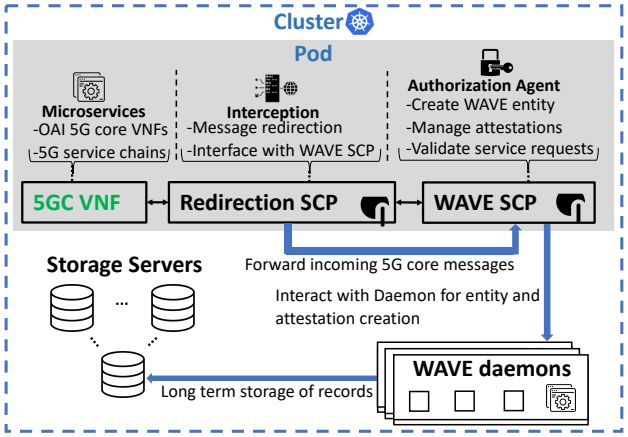

Fig. 4: Overview of the 5G-WAVE integrated platform to achieve decentralized inter-VNF authorization in the 5G core

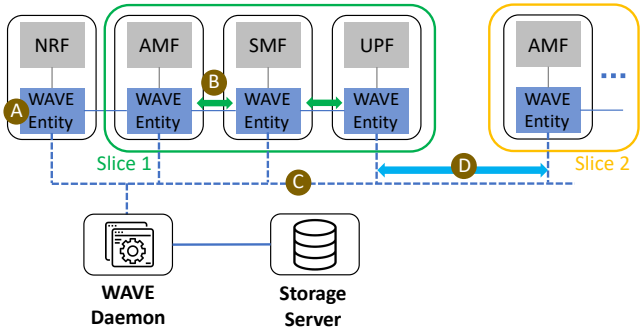

Fig. 5: Overview of intra/inter network slice interactions for the integrated 5G-WAVE platform

the database where the WAVE daemon publishes entities and attestations. Both WAVE daemons and storage servers can be distributed across multiple logical instances.

B. 5G-WAVE Integrated Network Slicing Overview

An overview of network slice interactions is illustrated in Figure 5. For simplicity, we have abstracted out the rSCP from the representation of individual pods. Two different network slices are shown with a single NRF and different AMF, SMF, UPFs. In Figure 5, (A)-(C) correspond to intra-slice flows while (D) is an inter-slice flow.

For the construction of a network slice in 5G-WAVE, the critical design element is the attachment of a WAVE entity to each 5G core VNF. Thus, for the flows in Figure 5, the first step in network slice construction is (A) the creation of the WAVE entity adjacent to each VNF. This occurs inside the wSCP of each pod in Figure 4. Hereafter, the VNFs will offload authorization functionality to these WAVE entities residing in each wSCP. During the network slice construction chain, each VNF will (B) authorize the subsequent VNF down the 5G core service chain. For example, for the 5G-AKA procedure, the involved VNFs (i.e., UDR, UDM, AUSF, AMF) will construct an authorization chain using their respective WAVE entities. The WAVE entities belonging to the VNFs of different network slices will be able to (C) communicate with each other using the same WAVE daemon.

In addition to intra-slice authorizations, (D) inter-slice authorizations can also be granted in the 5G-WAVE integrated platform. This facilitates the communication of VNFs from different network slices without relying on an NRF to distribute access tokens. For network slice handover scenarios of the 5G clients, inter-slice authorization chains can be created to secure the VNF communications.

C. Authorization Chain Message Exchange

To illustrate how the message flow within the OAI 5G core network is affected, a sample HTTP interaction between the NRF, AMF, and SMF is shown in Figure 6. For brevity, the rSCP has been abstracted out. Nevertheless, all the incoming messages to a given pod are still intercepted and forwarded through the rSCP. The individual HTTP interactions in this sample scenario have been labeled as belonging to either a 5G core service chain or the internal WAVE authorization flow.

The first set of messages occur during the instantiation of the 5G core VNFs, before they become service-ready. During this pre-instantiation phase, the WAVE flow (1.1)-(1.4) is repeated for all the 5G core VNFs. In Figure 6, we show it occurring only for the AMF to simplify the illustration. First, (1.1) the WAVE entity is created inside the wSCP and (1.2) published through the WAVE daemon and storage server. The wSCP of the AMF (1.3) receives the necessary WAVE credentials related to its WAVE entity. After the entity creation, it will (1.4) create attestations for service consumers that will be contacting it with service requests.

At this point, the VNF is service-ready and the relevant 5G core service chains can begin. In Figure 6, we illustrate the VNF discovery service request from AMF. The AMF will (2.1) contact the NRF, seeking a target SMF to set up the network slice. This message is intercepted by the pod sandbox of the NRF so that the authorization can be (2.2) verified through the WAVE flow. Then, the original message is forwarded to the NRF where (2.3) the profile of a candidate SMF is retrieved and (2.4) sent back to the AMF. The same process repeats itself for any subsequent 5G core service chains.

D. wSCP and rSCP Message Flows

The detailed message flow between two sample 5G core VNFs is shown in Figure 7. Each pod is made up of the 5G core VNF as well as the wSCP and rSCP. Before the beginning of this message exchange, each wSCP has created its own WAVE entity as described in Section III-A.

Upon instantiation, ❶ VNF-1 will send a list of target service providers to its adjacent wSCP. This will be a list of IP addresses identifying the target VNFs from which VNF-1 will seek to consume a service. For the depiction in Figure 7, VNF-1 is seeking this authorization from VNF-2. Hence, after wSCP-1 receives the target list of service providers (i.e., VNF-2), it will contact them ❷ with an authorization request. This request will be sent with the hash of the WAVE entity created by wSCP-1. In response, ❸ wSCP-2 will send back the hash

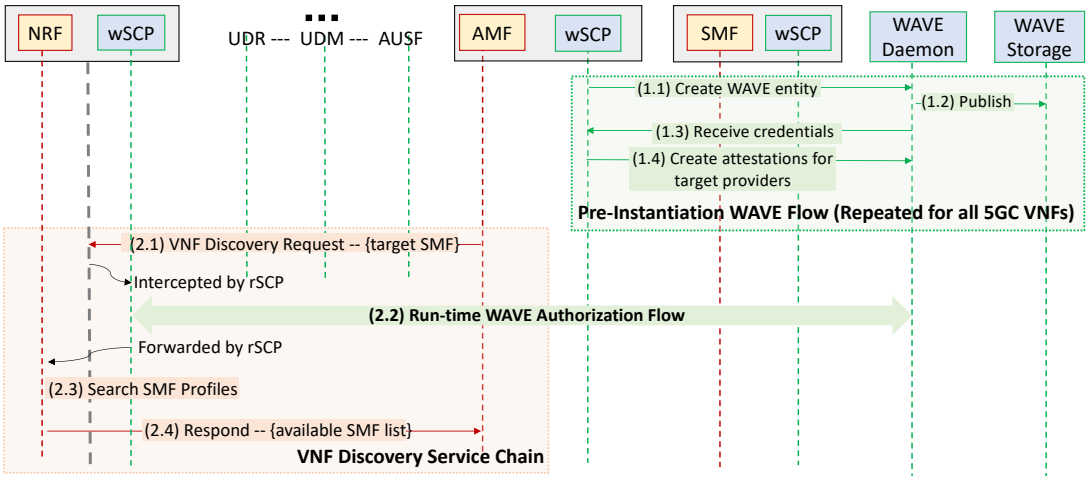

Fig. 6: Overview of the new message flow between WAVE entities and 5G core VNFs for authorization during network slice construction. The rSCP has been abstracted out from the pods for simplicity.

of the attestation and hash belonging to its own WAVE entity. This acknowledges that VNF-1 has been authorized by VNF-2.

To initiate the relevant 5G service chain, ❹ VNF-1 sends a service consumption request to VNF-2. Upon entering the pod network sandbox, this message is redirected to rSCP-2 instead of directly going to VNF-2. This redirection is achieved by performing port forwarding on the incoming HTTP requests so that they end up at the rSCP, which in turn communicates with the wSCP. Inside rSCP-2, the HTTP transactions are parsed and the source IP address (i.e., the IP of VNF-1) is extracted. The rSCP-2 ❺ forwards the IP of the sender to wSCP-2 where it triggers the verification process within the WAVE authorization chain. If the authorization request is successfully verified, the wSCP-2 ❻ responds with a valid verification message to rSCP-2 which ❼ forwards the original HTTP request from ❹ to VNF-2. However, if the authorization response from WAVE is not verified, the service request gets denied and is not forwarded to the 5G component. Once VNF-2 receives the HTTP message, it ❽ responds to VNF-1.

E. Implementation Details

We use Kubernetes for large-scale deployment of the 5G core VNFs with WAVE entities across multiple Virtual Machines (VMs). The WAVE daemon and storage server are also containerized pods, that have been deployed locally. Even though the storage server can be hosted on the cloud, we opt for a local deployment to reduce the latency of communication between the storage server and WAVE daemon.

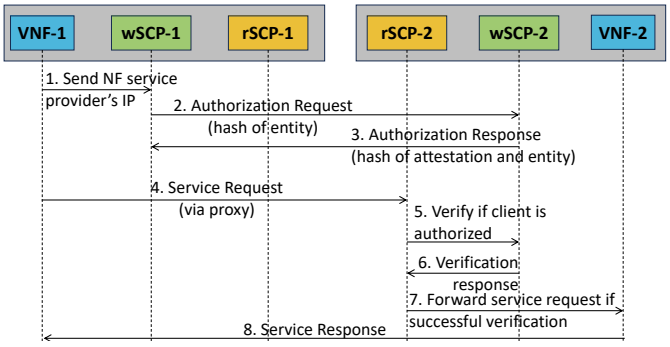

Fig. 7: Detailed message flow between the rSCP and wSCP of two 5G core VNFs

5G core VNFs. We use v1.2.1 of the OAI 5G core [7] VNFs for implementation. We also introduce an additional code block into each VNF's instantiation to trigger the authorization request for attestation from target providers wSCP.

Design of wSCPs. We design the wSCPs from scratch as independent HTTP servers built using C++17. In addition to the HTTP server, the wSCP container is also populated with the relevant WAVE scripts for the creation of WAVE entities and attestation. We use WAVE binaries *waved* and *wv* from the WAVE open-source repository [15] as WAVE daemon and agent to create WAVE objects in each wSCP. The scripts are executed at the relevant endpoints of the wSCP HTTP server.

Design of rSCPs. The rSCP is another HTTP server. However, to facilitate the manipulation of the HTTP messages, we use Python to implement the rSCP. We chose to separate the rSCP and wSCP in this manner to further decentralize the deployment based on microservice design principles. This allows for a more flexible production-grade deployment where the different sub-processes of the 5G-WAVE integrated platform are executed by independent entities.

Pod sandbox initContainer. We modify the IPTABLES of the pod using an initContainer so that incoming/outgoing messages to/from the VNFs are redirected to the rSCPs.

IV. Experimental Setup

Our experimental setup is shown in Figure 8. The infrastructure is comprised of two physical nodes that are hosting an OpenStack orchestrator. One of the nodes is the controller, networking and block storage node while the other serves as a compute node with the KVM hypervisor. The physical nodes used are two Precision 7920 Tower servers with 2 x Intel Xeon Gold 5218R 2.1GHz CPUs, 512GB RAM, 1TB disk space, and running Ubuntu 20.04.

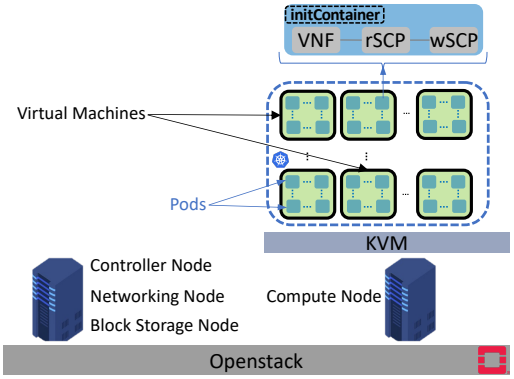

Fig. 8: Testbed with Openstack virtual machines hosting a Kubernetes cluster

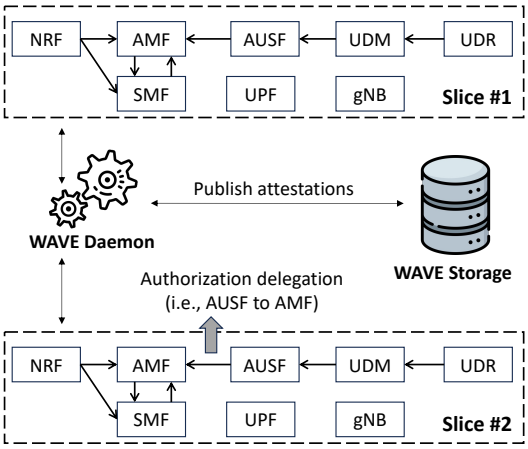

Fig. 9: Authorization chains in two different network slices

The compute node is running a total of 14 VMs, which are hosting a Highly Available (HA) Kubernetes cluster for production-grade testing. The roles and the flavors of the VMs in the cluster are given in Table I.

Our 5G core deployment consists of 6 control plane VNFs (i.e. NRF, UDR, UDM, AUSF, AMF, SMF) and 1 user plane VNF (i.e., UPF) in each slice deployed as Kubernetes pods. Each pods consists of 3 containers: 5G core VNF, wSCP and rSCP. We use the v1.2.1 OAI implementation of the 5G core with gNBSIM [8] entity for the RAN.

V. Performance Evaluation

We conduct two sets of experiments to evaluate the performance of 5G-WAVE. In particular, we measure the HTTP request-response times from two perspectives: 1) authorization overhead in single slice and 2) scalability in multiple slice deployment. For the first set of experiments, we deploy a single slice of the 6 core VNFs in the control plane along with a UPF and a gNBSIM. For the second set of experiments, we increase the number of network slices to measure the authorization overhead with more VNFs and analyze how 5G-WAVE scales with network size. Within each network slice, there is a VNF authorization chain. For scalability tests, we deploy multiple logically-isolated slices sequentially for use cases with strict security requirements. Two such slices are illustrated in Figure 9. We perform 20 and 10 iterations, respectively, for the two experiments to average our results.

A. Single slice authorization overhead

So far, none of the open-source 5G core implementations have OAuth 2.0 authorization functionality built into them. Thus, the performance overhead of 5G-WAVE is compared against a 5G deployment with no built-in authorization.

While analyzing the message flows between VNFs, we identify 7 unique pairs of requester-authorizer VNFs for attestation creation. The mean and standard deviation of the time taken to create attestations for each pair are given in Table II. The client VNF forms the requesting entity in WAVE and the host VNF forms the authorizing entity. We further identify 14 distinct API requests for a functional deployment of the 5G core, which are listed in Table III. Each request involves a WAVE verification operation by the authorizing entity. We present the duration of a single verification operation in Table III and the total overhead for the HTTP request in Figure 10.

In Figure 10, two sets of measurements are stacked for each bar representing 5G-WAVE. The results show the time cost for 1) the 'verification' operation of WAVE (stacked on top) and 2) 5G core service chain and rerouting overhead via rSCP (stacked on bottom). The majority of the overhead is introduced from the WAVE verification flows. The transactions in the 5G core can be categorized into three groups: 1) VNF instance registration (bars 1-3), 2) 5G-AKA service chain (bars 4-11), and 3) Packet Data Unit (PDU) session setup (bars 12-14). The WAVE verification bars are split further for the 5G-AKA operations to illustrate the number of verification operations in each API request and the time cost for them.

VNF registration. During instance registration, both AMF and SMF send requests to NRF. We see from Table III, that the WAVE verification process takes almost the same amount of time for the three HTTP requests (≈ 113 ms). Looking at Figure 10, these correspond to the PUT requests going from the AMF/SMF to the NRF and incur the lowest overhead among all the HTTP transactions.

TABLE I: HA Kubernetes cluster VM flavors

Node	Instances	vCPUs	RAM (GB)	Disk (GB)
Control	2	2	8	40
Worker	11	4	8	80
Storage	1	4	8	80

TABLE II: Time taken (in ms) by authorizer VNFs to create attestations for requesting client VNFs

Requester (Client)	Authorizer (Host)	Mean	Std
AMF	NRF	156.52	12.35
SMF	NRF	133.16	17.54
UDM	UDR	159.40	28.64
AUSF	UDM	197.80	21.72
AMF	AUSF	184.90	18.26
SMF	AMF	195.90	21.98
AMF	SMF	183.94	15.94

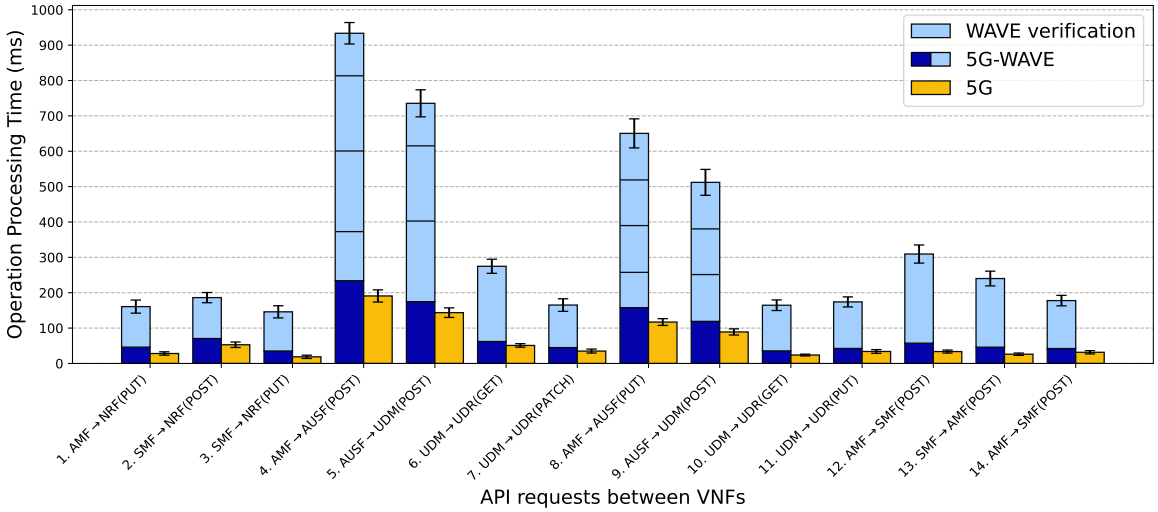

Fig. 10: Latency overhead comparison of the 5G-WAVE integrated platform with the OAI 5G core deployment.

5G-AKA service chain. This is the primary UE authentication service chain that comprises a series of API requests between the AMF, AUSF, UDM, and UDR. This service chain is triggered when the UE sends a registration request to the AMF. The AMF interacts with the AUSF which in turn requests the UDM to generate the authentication vector, which further propagates it to the UDR. This is a service chain of three VNFs that involves 4 sequential API transactions for the original UE authentication request to be completed.

TABLE III: Time taken (ms) for verification of WAVE authorization for different Client→Host pairs, measured on the authorizing entity's end. The Client is the requesting WAVE entity and the Host is the authorizing WAVE entity.

	Client	Host	Service Request	Time
VNF instance registration				
1	AMF	NRF	PUT: nnrf-nfm (NF registration)	114.55
2	SMF	NRF	POST: nnrf-nfm (Subscribe to updates)	115.65
3	SMF	NRF	PUT: nnrf-nfm (NF registration)	110.75
Authentication and Key Agreement				
4	AMF	AUSF	POST: nausf-auth (UE authentication)	138.80
5	AUSF	UDM	POST: nudm-ueau (Generate authentication vectors)	227.95
6	UDM	UDR	GET: nudr-dr (Query UE authentication data)	212.65
7	UDM	UDR	PATCH: nudr-dr (Update UE authentication data)	120.25
8	AMF	AUSF	PUT: nausf-auth (Authentication confirmation)	99.90
9	AUSF	UDM	POST: nudm-ueau (Inform about authentication result)	132.30
10	UDM	UDR	GET: nudr-dr (Query UE authentication data)	129.05
11	UDM	UDR	PUT: nudr-dr (Update UE authentication data)	131.60
PDU Session Setup				
12	AMF	SMF	POST: nsmf-pdusession (Create SM Context)	252.00
13	SMF	AMF	POST: namf-comm (N1-N2 message transfer)	194.10
14	AMF	SMF	POST: nsmf-pdusession (Update SM Context)	135.55
			Mean Time	151.08

Thus, we can see in Figure 10, in the `AMF→AUSF(POST)` and `AMF→AUSF(PUT)` operations, the WAVE verification duration is the sum of four verification operations. Although the UE authentication has `4x` verification overhead, it is a one-time process and the security benefit of WAVE outweighs the overhead in total request completion time.

PDU session setup. PDU session setup consists of two `AMF→SMF(POST)` API calls. Compared to the 5G-AKA service chain, the WAVE overhead is smaller for PDU session setup. A single UE can instantiate multiple PDU sessions within the same network slice. This overhead will not introduce significant delays in the 5G core operations.

B. Scalability overhead with multiple slices

To evaluate how the 5G-WAVE framework scales in a larger network, we deploy multiple slices which are isolated from each other. After VNF instance registration, SMF sends a heart-beat timer request to NRF periodically. This is a `PATCH` HTTP request that involves a single WAVE verification operation. We measure the average time taken for this request to complete as we keep increasing the number of slices in the deployment. This gives us the estimate of how long 5G-WAVE will take for a single HTTP message. It can be observed from Figure 11 that for a single slice deployment, the overhead by 5G-WAVE is 155ms. We further observe a linear trend in the increase of request completion time with the number of slices. This is expected because there is an increase in computational demand from the WAVE daemon with more slices. Note that even when the number of slices grows by 10x, the time taken only increases by 1.4x. Thus, we can claim that 5G-WAVE scales well with the network size.

VI. Security Analysis

In this section, we describe our two-stage security analysis of the 5G-WAVE integrated platform. Firstly, we discuss how specific OAuth attack vectors are addressed in Table IV. Secondly, to illustrate how the proposed design improves the

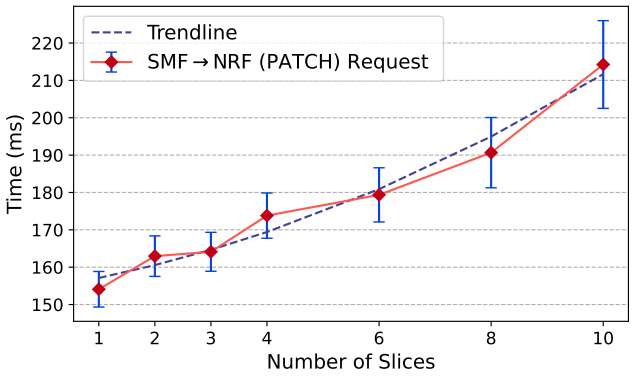

Fig. 11: Latency overhead with increasing number of slices

flexibility and security of the 5G core, we reference Key Issues (KIs) from the 3GPP standardization in Table V.

A. Comparing 5G-OAuth and 5G-WAVE

To ensure inter-slice and intra-slice security, 5G core SBA must adopt a Zero-Trust Architecture (ZTA) [16], which demands all network entities to be authenticated, authorized, and continuously verified before being granted access to resources. 3GPP has standardized OAuth 2.0 as an authorization framework for VNF service access to achieve SBA domain security. Within this framework, the NRF acts as an authorization server that distributes access tokens to service consumer VNFs [2], [3]. Either directly or as a result of misconfiguration, OAuth 2.0 is vulnerable to the attack types listed in Table IV.

When OAuth 2.0 is replaced with a decentralized approach, we remove the dependency on the NRF as a central authorization server. Each VNF service provider becomes both the authorization and resource server for their own services. To realize this idea, our decentralized authorization framework of choice is WAVE. It has been shown that WAVE performs better than traditional authorization frameworks by providing stronger security guarantees with comparable performance [4].

In WAVE, services are granted in the form of encrypted attestations rather than tokens. As a result, the resource requester and resource provider are bound to each other in the WAVE authorization chain without involving a third-party server. These attestations can be stored in a distributed manner using horizontally scalable servers without compromising system security. The critical security features provided by incorporating WAVE into the 5G core are listed below.

TABLE IV: Comparing the detection and mitigation capabilities of the standardized 5G-OAuth framework with the 5G-WAVE integrated platform (✓: addressed, ▲: addressed with additional client-side configuration, x: vulnerable)

Attack Type	5G-OAuth	5G-WAVE
Authorization Code Injection	▲	✓
Access Token Leakage	▲	✓
Credential Phishing Attacks	▲	✓
Authorization Flooding	x	✓
Access Token Hijacking	▲	✓

TABLE V: Summary of 3GPP key issues addressed by the 5G-WAVE integrated platform (✓: directly addressed, ▲: addressed with proper configuration, ⧖: future work)

3GPP #	KI #	Description	5G-WAVE
33.813 [20]	5	token handling between slices	✓
33.886 [21]	2	temporary slice authorization	▲
33.874 [22]	3	AF authorization	▲
33.813 [20]	1	slice-specific authorization	⧖

- Confidentiality, integrity, and replay protection are supported for inter-VNF communication.
- The 5G core network slice topology in each administrative domain is maintained separately. Thus, for Multi-Operator Networks (MNOs), the VNF topologies are hidden from entities in different trust domains.
- Instead of relying on an authorization server, the service request messages will be validated by the service provider VNFs. Invalid or unauthorized messages will be rejected or discarded according to the protocol specification.

A comparison of the 5G-OAuth and 5G-WAVE is provided in Table IV. The specific attack vectors are explained below.

As a Man-in-the-Middle (MiTM) attacker, adversaries can perform authorization code injections, access token hijacking and take advantage of token leakages in the OAuth 2.0 framework [17]–[19]. Without additional client-side security implemented on the VNFs (e.g., URL redirects, TLS), the 5G core remains vulnerable to these OAuth 2.0 attack vectors. On the other hand, WAVE prevents these attacks by removing the reliance on access tokens. By extension, this no longer requires the employment of risk-inducing information elements such as authorization codes and access scopes.

As a centralized authorization server, the NRF is vulnerable to authorization flooding attacks. Adversaries can execute a Distributed Denial of Service (DDoS) by generating a high volume of access token requests directed at the NRF. With the decentralization provided by the 5G-WAVE integration, this vulnerability (i.e., NRF as OAuth 2.0 server) is removed.

Adversaries using phishing attacks will impersonate a legitimate service requester to obtain illegal access tokens and/or access credentials from legitimate OAuth 2.0 entities. However, WAVE uses custom entities adjacent to each 5G core VNF to ensure their authenticity. As a result, any impersonator not accompanied by a WAVE entity will not be able to infiltrate the authorization chain of the 5G-WAVE integrated platform.

Several attacks from Table IV can be addressed by additional client-side configuration. However, this requires additional policies, development effort, and expertise. With the 5G-WAVE integrated platform, all the attack types are addressed out of the box. This enables mobile operators to focus on deployment without concern about the underlying security.

B. Addressing 3GPP Key Issues

[33.813 KI 5] The NRF can be deployed at different administrative hierarchies (e.g., Public Land Mobile Network level, shared-slice level, slice-specific level). The tokens distributed by the NRF can be used to access services of the same type of

VNF service providers across multiple slices. This can lead to illegal access requests since different network slices may have alternative service consumption policies for their VNF types. The 5G-WAVE integrated platform directly addresses this KI as it removes the need for tokens altogether.

[33.886 KI 2] For the graceful termination of data sessions, UEs are allowed access to temporary slices. This requires efficient authorization mechanisms to moderate access to such slices. By creating a joint authorization chain between the UE's current and target temporary slice, authorized access can be provided with the 5G-WAVE integrated platform.

[33.874 KI 3] External Application Functions (AFs) [23] can be exposed to the 5G core SBA for carrying out custom operator functions. Without proper authentication and authorization, malicious AFs can gain access to sensitive information by accessing network slices. The operator can attach the AFs to the 5G-WAVE integrated authorization chain ensuring that the service provider VNFs can directly provide authorization to AFs for service consumption.

[33.813 KI 1] Network slice-specific authentication and authorization has become a fundamental requirement. To address this KI, 3GPP has already standardized the Network Slice-Specific Authentication and Authorization Function (NSSAAF) [24]. This additional VNF will serve as an independent authentication and authorization server for moderating access to individual slices. Currently, the 5G-WAVE integrated platform cannot directly be used to address this KI. As a future work, we plan to store slice-specific metadata inside the wSCPs to keep track of the VNF authorizations along with their network slices.

VII. Related Work

With the growing popularity of open-source 5G, various testbed-based studies have spawned investigating several scalability aspects [25], [26]. Furthermore, as the microservice model of deployment for applications has gained more traction, authorization methodologies have adapted to accommodate the growing volume of interactions.

Focusing on the security aspects of 5G, Edris et al. [27] proposed a single sign-on federated identity for allowing multiple users to access services offered by providers. Similarly, a higher level authorization is proposed in [28], which focuses on handling authorization for a multi-tenant and operator ecosystem with a distributed and virtual authorization agent. These approaches do not investigate the authorization between individual VNFs but rather focus on the authorization of users and service providers at a higher level.

To defend against the reusability of access tokens, Zhang et al. [29] proposed to use a trusted NRF as a certificate authority to further vouch for the access tokens with a certificate. While the approach adds another layer of security, it further centralizes the authorization mechanism of the 5G architecture. Furthermore, token hijacking and interception attack vectors, as well as authorization flooding attacks, are not addressed.

A different approach is presented in [30] for network slice specific authorization. Access control for specific slices is offloaded to the 3rd party UE operators, instead of relying on the 5G core network. Their work is orthogonal to ours as they focus on UE authentication and access control (AAC) and design additional RAN NFs to delegate control to the 3rd parties. By contrast, we address mutual authorization between VNFs for service consumption.

Akon et al. [31] conducted a formal analysis of OAuth 2.0-based access control mechanism in the 5G core as outlined in the 3GPP specifications, and they identified 5 major vulnerabilities in the design. This work further strengthens our motivation to remove a central authorization server in the 5G core. A very brief discussion of distributed authorization in the 5G micro-service architecture is offered in [32], but it lacks implementation and evaluation details.

Looking into 5G and beyond, Atalay et al. [33] proposed an authorization mechanism for the the O-RAN ecosystem [34]. Their work similarly relies on SCPs for offloading functional security from the main applications. Therefore, our proposed 5G-WAVE integrated platform can be easily applied to their design for enhanced security features.

To the best of our knowledge, no existing work explores application of entity-based decentralized authorization in the 5G network architecture. Our work is the first to present the integration of decentralized authorization into the 5G core.

VIII. Conclusion

5G networks have already been deployed globally, and a continued reliance on a central authorization server in 5G core will lead to security and performance issues. To address this, we introduced the 5G-WAVE integrated platform for decentralizing inter-VNF communications. Our design utilizes an interception/authorization pipeline based on SCPs attached to individual 5G core VNFs deployed as Kubernetes pods. We evaluated the overhead of our proposed solution in comparison with native 5G deployments without any authorization. The 5G-WAVE integrated platform introduces 155ms authorization latency overhead to HTTP transactions in the 5G core service chains for a single slice. Additionally, the authorization latency increases linearly by 1.4 times with a 10-fold growth in the network. The 5G-WAVE framework enhances security by addressing OAuth 2.0 vulnerabilities and 3GPP key issues for network slicing security. The authors have provided public access to their code at [35].

Acknowledgment

We thank the anonymous reviewers for their valuable feedback. This work was supported in part by the U.S. Defense Advanced Research Projects Agency (DARPA) under agreement number HR001120C0155. The views, opinions, and findings contained in this article are those of the authors and should not be interpreted as representing the official views or policies, either expressed or implied, of the Defense Advanced Research Projects Agency or the Department of Defense.

REFERENCES

[1] 3GPP, "System Architecture for the 5G System (5GS)," 3rd Generation Partnership Project (3GPP), TS 23.501 V17.8.0 , Mar. 2023.

[2] ——, "Security Architecture and Procedures for 5G System," 3rd Generation Partnership Project (3GPP), TS 33.501 V17.9.0 , Mar. 2023.

[3] ——, "5G System; Network Function Repository Services; Stage 3," 3rd Generation Partnership Project (3GPP), TS 29.510 V17.9.0 , Mar. 2023.

[4] M. P. Andersen, S. Kumar, M. AbdelBaky, G. Fierro, J. Kolb, H. S. Kim, D. E. Culler, and R. A. Popa, "WAVE: A Decentralized Authorization Framework with Transitive Delegation," in *Proceedings of the 28th USENIX Security Symposium*. Santa Clara, CA: USENIX Association, Aug. 2019, pp. 1375–1392.

[5] Microsoft, "Sidecar Pattern - Azure Architecture Center," Jun. 2023. [Online]. Available: https://docs.microsoft.com/en-us/azure/architecture/patterns/sidecar

[6] 3GPP, "5G System; Technical Realization of Service Based Architecture; Stage 3," 3rd Generation Partnership Project (3GPP), TS 29.500 V17.10.0 , Mar. 2023.

[7] "OpenAirInterface 5G Software Alliance for Democratising Wireless Innovation," https://openairinterface.org/.

[8] Rohan, "Gnbsim: Ue and gnb simulator," https://gitlab.eurecom.fr/kharade/gnbsim, 2023.

[9] M. P. Andersen, J. Kolb, K. Chen, G. Fierro, D. E. Culler, and R. Katz, "Democratizing Authority in the Built Environment," *ACM Transactions on Sensor Networks*, vol. 14, no. 3-4, dec 2018.

[10] M. P. Andersen, "Decentralized Authorization with Private Delegation," Ph.D. dissertation, 2019. [Online]. Available: https://www2.eecs.berkeley.edu/Pubs/TechRpts/2019/EECS-2019-113.pdf

[11] M. P. Andersen, J. Kolb, K. Chen, G. Fierro, D. E. Culler, and R. A. Popa, "WAVE: A Decentralized Authorization System for IoT via Blockchain Smart Contracts," *University of California at Berkeley, Tech. Rep*, 2017.

[12] R. Rivest and B. Lampson, "Sdsi – a simple distributed security infrastructure," *See the SDSI web page at http://theory.lcs.mit.edu/ cis/sdsi.html*, 08 1996.

[13] A. Birgisson, J. G. Politz, Ú. Erlingsson, A. Taly, M. Vrable, and M. Lentczner, "Macaroons: Cookies with contextual caveats for decentralized authorization in the cloud," in *21st Annual Network and Distributed System Security Symposium, NDSS 2014, San Diego, California, USA, February 23-26, 2014*. The Internet Society, 2014.

[14] OpenAirInterface, "CN5G - gitlab," https://gitlab.eurecom.fr/oai/cn5g/oai-cn5g-fed, 2023, (Accessed on 05/31/2023).

[15] M. Anderson, "WAVE Go," Oct. 2022. [Online]. Available: https://github.com/immesys/wave

[16] S. Rose, O. Borchert, S. Mitchell, and S. Connelly, "Zero Trust Architecture," *CSRC | NIST*, Aug. 2020.

[17] R. 6819, "OAuth 2.0 Threat Model and Security Considerations," https://datatracker.ietf.org/doc/html/rfc6819, Jan 2013.

[18] IETF, "OAuth-Security-Topics-22," https://datatracker.ietf.org/doc/html/draft-ietf-oauth-security-topics-16#section-2.4.2, Mar 2023.

[19] R. Yang, G. Li, W. C. Lau, K. Zhang, and P. Hu, "Model-based Security Testing: An Empirical Study on OAuth 2.0 Implementations," in *Proceedings of the 11th ACM on Asia Conference on Computer and Communications Security*, 2016, pp. 651–662.

[20] 3GPP, "Study on Security Aspects of Network Slicing Enhancement," 3rd Generation Partnership Project (3GPP), TR 33.813 V16.0.0, Jul. 2020.

[21] ——, "Study on enhanced security for Network Slicing Phase 3," 3rd Generation Partnership Project (3GPP), TR 33.886 V0.4.0, Feb. 2023.

[22] ——, "Study on enhanced security for network slicing phase 2," 3rd Generation Partnership Project (3GPP), TR 33.874 V18.1.0, Sep. 2022.

[23] ——, "5G System; Application Function Event Exposure Service; Stage 3," 3rd Generation Partnership Project (3GPP), TS 29.517 V18.1.0 , Mar. 2023.

[24] ——, "Network Slice-Specific and SNPN Authentication and Authorization; Stage 3 services;," 3rd Generation Partnership Project (3GPP), TS 29.526 V18.1.0, Mar. 2023.

[25] T. O. Atalay, D. Stojadinovic, A. Famili, A. Stavrou, and H. Wang, "Network-Slice-as-a-Service Deployment Cost Assessment in an End-to-End 5G Testbed," in *IEEE Global Communications Conference (GLOBECOM)*, 2022, pp. 2056–2061.

[26] T. O. Atalay, D. Stojadinovic, A. Stavrou, and H. Wang, "Scaling Network Slices with a 5G Testbed: A Resource Consumption Study," in *2022 IEEE Wireless Communications and Networking Conference (WCNC)*, 2022, pp. 2649–2654.

[27] E. K. K. Edris, M. Aiash, and J. K.-K. Loo, "Network Service Federated Identity (NS-FId) Protocol for Service Authorization in 5G network," in *2020 Fifth International Conference on Fog and Mobile Edge Computing (FMEC)*. IEEE, 2020, pp. 128–135.

[28] S. Wong, N. Sastry, O. Holland, V. Friderikos, M. Dohler, and H. Aghvami, "Virtualized Authentication, Authorization and Accounting (V-AAA) in 5G Networks," in *2017 IEEE Conference on Standards for Communications and Networking (CSCN)*. IEEE, 2017, pp. 175–180.

[29] Y. Zhang, C. Liu, S. Liu, and F. Pan, "SETOKEN: A Secure Protection Mechanism based on Service API for 5G Network Access Token," in *2021 2nd International Conference on Electronics, Communications and Information Technology (CECIT)*. IEEE, 2021, pp. 1139–1143.

[30] S. Behrad, E. Bertin, S. Tuffin, and N. Crespi, "5G-SSAAC: Slice-Specific Authentication and Access Control in 5G," in *2019 IEEE Conference on Network Softwarization (NetSoft)*. IEEE, 2019, pp. 281–285.

[31] M. Akon, T. Yang, Y. Dong, and S. R. Hussain, "Formal Analysis of Access Control Mechanism of 5G Core Network," in *Proceedings of the 2023 ACM SIGSAC Conference on Computer and Communications Security*, ser. CCS '23. New York, NY, USA: Association for Computing Machinery, 2023, p. 666–680. [Online]. Available: https://doi.org/10.1145/3576915.3623113

[32] D. Guija and M. S. Siddiqui, "Identity and Access Control for Micro-Services Based 5G NFV Platforms," in *Proceedings of the 13th International Conference on Availability, Reliability and Security*, ser. ARES '18. New York, NY, USA: Association for Computing Machinery, 2018. [Online]. Available: https://doi.org/10.1145/3230833.3233255

[33] T. O. Atalay, S. Maitra, D. Stojadinovic, A. Stavrou, and H. Wang, "Securing 5G OpenRAN with a Scalable Authorization Framework for xApps," in *IEEE Conference on Computer Communications (INFOCOM)*, 2023, pp. 1–10.

[34] L. Foundation, "O-RAN ALLIANCE e.V," https://www.o-ran.org/.

[35] "5G-WAVE-Infocom2024," Jan. 2024. [Online]. Available: https://github.com/pragyasharmaa/5G-WAVE-Infocom2024