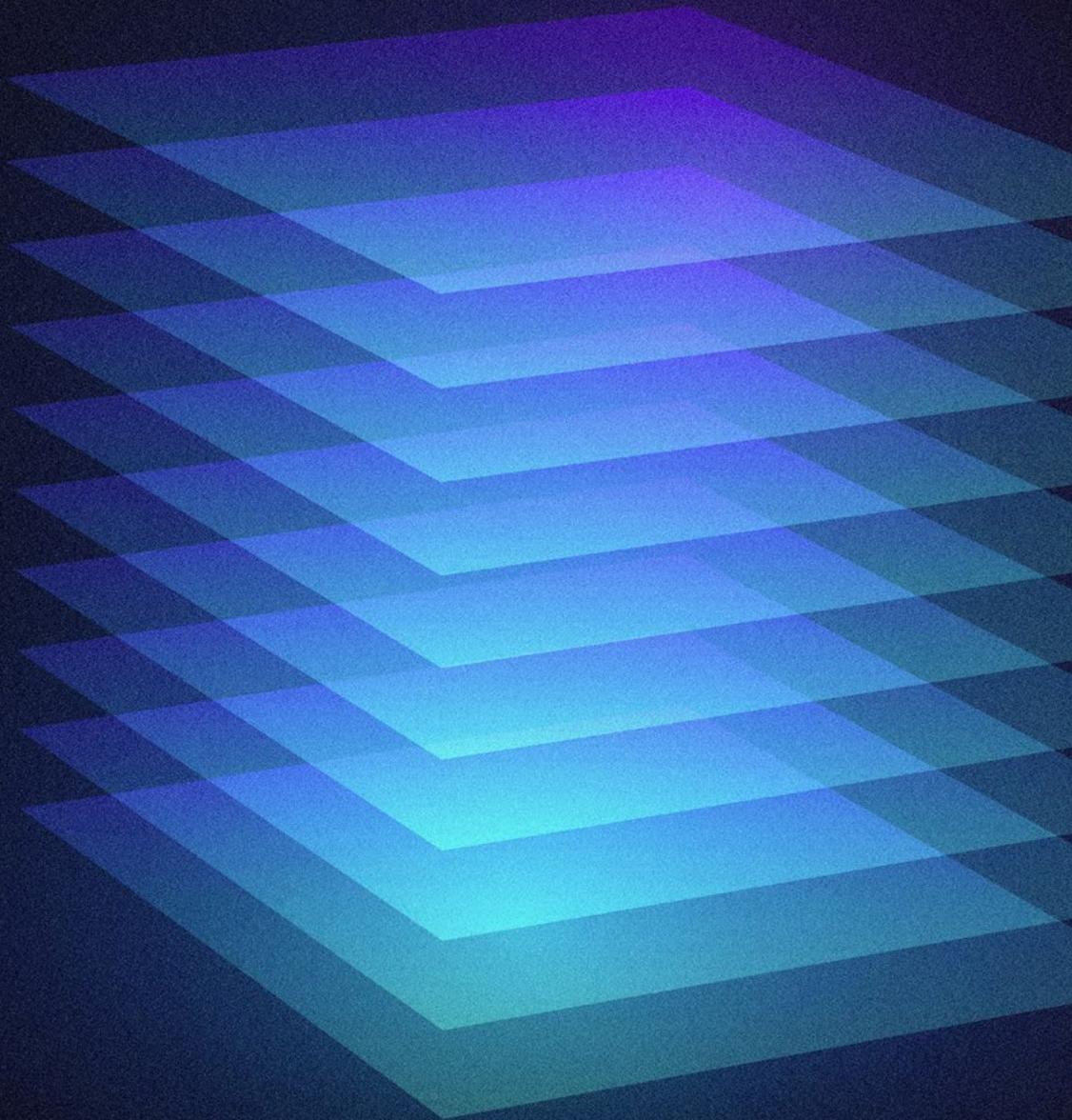




AdaptiveMobile Security

WHITE PAPER

# A Slice in Time: Slicing Security in 5G Core Networks





## Table of Contents

	Executive Summary	3
1.	Introduction	4
2.	5G Architecture	5
2.1	Service Based Architecture and Slicing	5
2.2	Slicing in the 5G Core Network	6
2.2.1	Security Zones as a Consequence of Slicing	7
2.3	Slice Types and Properties	8
2.3.1	Slice Types	8
2.3.2	Slice Properties and Security Attributes	9
2.4	Authorization in 5G Service Based Architecture	10
2.4.1	Network Repository Function (NRF)	10
2.4.2	Service Communication Proxy (SCP)	15
2.5	User Equipment Access to a Slice	16
2.5.1	Simple Slice Access	16
2.5.2	Slice Specific Authentication	18
3.	Security Challenges for 5G Slicing and Countermeasures	19
3.1	Security Challenges	19
3.1.1	Security Challenges Through Legacy Interworking	19
3.1.2	Security Challenges Through Increased Complexity	20
3.1.3	Security Challenges Through Configuration Mistakes and Missing Layer Matching	22
3.1.3.1	Vulnerability: Malicious Access to Different Slice by Modifying Slice Differentiator	22
3.1.3.2	Vulnerability: Specifying Different Slice in HTTP Header for DoS	24
3.1.4	Security Challenges Through Missing Security Zones	26
3.1.5	Vulnerability: Security Challenges Through Missing Security Granularity	27
3.2	Security Countermeasures and New Features	29
4.	Summary and Conclusion	32
	Abbreviations	33
	References	34
	Annex: Timeline/Responsible Disclosure Process	35



## Executive Summary

The world can't help but change around us, and this will become apparent once again with the rollout of 5G.

5G has the potential to revolutionize society and dramatically alter our everyday lives. From nation states to individual citizens, we will all be impacted. Industries will change and new business models will be created, as mobile telecoms are transformed. With this change, exciting new use cases will emerge across many verticals such as Automotive, Healthcare and Entertainment.

Two fundamental technologies will be at the heart of this 5G transformation: Slicing and Virtualization. Although not yet widely deployed, their use will accelerate as 5G networks are built-out around the globe.

In anticipation of this rapid 5G expansion, AdaptiveMobile Security has undertaken an in-depth analysis of the security of these two bleeding edge technologies, Slicing and Virtualization.

This analysis is of immediate and paramount importance as 5G networks will encompass many new partners and an array of new interfaces, and so it should be expected that malicious actors will attempt new types of attacks.

The question is whether the mechanisms, currently defined in the standards, will be sufficient to stop these attackers and defend against their threats.

The answer is concerning, as three major attack scenarios were uncovered, which cannot be mitigated with today's approach and technologies:

- User data extraction - in particular location tracking
- Denial of service against another network function
- Access to a network function and related information of another vertical

AdaptiveMobile Security has submitted these vulnerabilities as a Common Vulnerability Disclosure (CVD) to the GSM Association (GSMA). The CVD has been accepted, with potential countermeasures proposed.

Moving forward, MNOs will need to be cognizant of these attack scenarios, as, be in no doubt, attackers will try and use them for nefarious purposes.



# 1. Introduction

New vertical use cases are the main driver for the deployment of 5G networks. It is anticipated that these use cases will depend on network slicing technology and network function virtualization. While the terms slicing, virtualization, cloud, edge computing and programmable network are often used in conjunction, the combination of these technologies offers many opportunities for mobile network operators (MNO) and service providers. It allows them to quickly onboard new services, scale and adjust existing services to their needs, and provides the bedrock for an open ecosystem where any kind of service can benefit from the features of the 5G network. Slicing is seen as the main 5G innovation that will provide all of this. However, there is a wide-scale lack of knowledge of what exactly slicing is and an even bigger lack of knowledge on whether it is secure.

In this paper, we intend to demystify slicing and explain with examples how it works in the 5G core network. We will then outline what security technologies exist in 5G to protect slicing on the signalling layer and what the potential gaps in this technology are.

Virtualization, cloud and slicing are technologies that run on different layers and different parts of the network. They complement each other and together enable the slicing concept for verticals. In this paper we will focus on the signalling layer. The signalling layer is the telecommunication specific application layer between network functions. The virtualization layer beneath may offer its own security features e.g. for authentication, communication security etc. but these are outside the scope of this paper.

This white paper is split into 3 main parts. First, in chapter 2 we will explain the existing 5G architecture and security features, focusing on:

- Introduction to the 5G architecture
- Concept of core network slicing
- The different slice types and properties
- Existing 5G security features that relate to slicing
- Authorization in the 5G Service Based Architecture (SBA) using Network Repository Function (NRF) and Service Communication Proxy (SCP)
- Security from a terminal and user perspective.

Next, in chapter 3, we look at the security challenges introduced by slicing and bringing external parties into the network, and we will cover aspects like legacy interworking and configuration risks introduced through the increase in complexity. Then we will take a closer look at details on the existing security structure and describe how an attacker may exploit missing checks and careless configurations to gain unauthorized access to information and resources. We will present three attacks to core network slicing security. Details of the attacks: how to gain access to resources of another slice, and how to perform a DoS attack on another slice, can be found in section 3.1.3. In section 3.1.5 we explain how to extract user specific information like location from another slice. We will explain how the existing features can be extended and new measures added to provide security zones in the network.

Finally, the conclusion in section 4 will summarize the key aspects of existing slicing security features, describing the gaps and the features needed to close them.



## 2. 5G Architecture

### 2.1 Service Based Architecture and Slicing

The 5G standards introduced a new architectural concept called the Service Based Architecture (SBA), (see figure 1), which allows fast and easy integration of new network services and brings telecommunication networks' approach to design closer to that of IT-networks. Using this architecture, Network Functions (NFs) can be virtualized and provide their services, using the common HTTP/2 Internet protocol and REST API based Service Based Interfaces (SBI), to other network functions or external parties' "verticals". There are different approaches to virtualization depending on the architectural and business arrangements in the network. This paper will focus on signalling security for core network slicing. We assume that the communication between network functions is authenticated, with confidentiality and integrity protected using TLS or IPSec protocols.

All 5G network functions of the core network are connected to the SBA. Incidentally, this architecture is backwards compatible - 4G nodes might be connected to the SBA through interworking functions<sup>[6]</sup> for example, for SMS over NAS or SMS over IP.

### Service Based Architecture

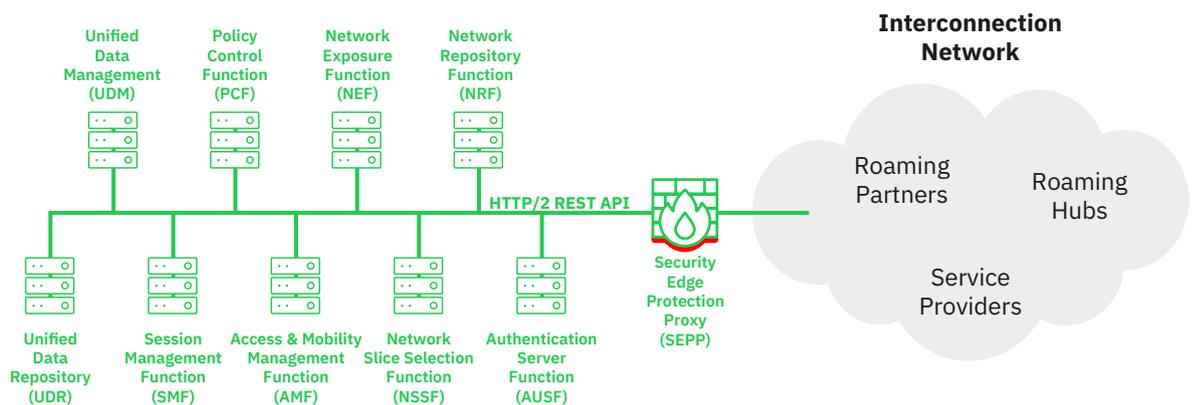


Figure 1: Service Based Architecture (SBA)

Each network function of the SBA can offer a service and so be a service producer, but every network function can also act as a service consumer when it requests a service from another network function. This way, in the SBA, each network function can be thought of as acting as a server or client. This is different from the classical HTTP client/server model.

5G has learned many security lessons from previous network generations and has taken measures to enhance its security. Examples of new security features in 5G are:

- Introduction of a Secure Edge Protection Proxy (SEPP) which protects the network from attacks arriving over the roaming network. The details are defined by the GSM Association (GSMA) in FS.36<sup>[1]</sup>
- Unified authentication framework for the various 5G access technologies and devices
- Protection of user privacy on the air interface
- Extended home network control for roaming users
- Introduction of the NRF authorization function into the network architecture

The SBA also offers security features to slices, which we will discuss in the next section.

<sup>[1]</sup> GSMA FS.36, "5G Interconnect Security", version 1.0, 20. November 2020

<sup>[6]</sup> Adaptive Mobile Security, Silke Holtmanns, Cathal McDaid, "Securing the Path from 4G to 5G: How to protect mobile networks and subscribers during migration", <https://www.adaptivemobile.com/downloads/securing-the-path-from-4g-to-5g>



## 2.2 Slicing in the 5G Core Network

The term slicing is used for the core network and the radio access network. In this paper we will focus on the core network part of the slice; it may have an extension into the Radio Access Network (RAN) but does not necessarily have to.

What is a slice? A slice is a logical block in the 5G core network to serve a specific purpose. This purpose can be massive machine type communication, mission-critical networks, private networks, streaming, automotive or a Mobile Virtual Network Operator (MVNO). For some of those purposes the slice details are specified. For other purposes, like MVNO, non-standardized slices can be defined by the hosting MNO. A slice has certain characteristics related to Quality of Service (QoS) e.g. supporting large bandwidths or short round-trip times etc. Depending on the specific purpose different slice types with specific QoS characteristics have been defined. For example, a streaming service and a private network of a hospital may need the same slice type with high throughput and low latency, even though the use cases are different.

A slice in the core network consists of a group of Network Functions (NFs) that support that slice. Those network functions can be exclusively assigned to that slice or be shared among different slices. The network functions can be virtual or physical. A physical node may host several network functions in 5G. A shared network function can provide services to several slices; the slices would not be separated on the transport and IP layer<sup>1</sup>. The management of slices, network functions, virtual machines, hypervisors etc. is a complex undertaking and will not be covered in depth here.

We assume, that some of these network functions might not belong to the hosting MNO, but a third party, and so a third party has access to the core network. Here's an example of this: a host MNO is hosting two slices, Slice 1 for gaming and Slice 2 for IoT devices. Slice 1 would have its own dedicated network functions like a Session Management Function (SMF) for session management, Policy Control Function (PCF) for policy control etc. But this Slice 1 would not only consist of the network functions in the Slice 1 "box", but also the UDR (Unified Data Repository), AMF (Access Management Function) and SEPP from the shared network functions "box". The network functions in the shared box would be available to be used by the slices and the hosting MNO.

Also, it is the same for Slice 2, it would consist of the functions of the Slice 2 "box" and additionally the shared functions UDR, SEPP and AMF would also be available to Slice 2.

A visual representation of the different "boxes" is shown below in Figure 2. This kind of approach to have not-shared, shared and slice-dedicated network functions is done to allow maximum flexibility. The Slice 1 gaming company may want to offer content close to the proximity of the user, so they take ownership of the UPF. On the other hand, the gaming company may not want to run the AMF themselves.

The hosting MNO wants to enlarge its business, but still not give third parties full access to its network functions, so it is not sharing some of the functions of the network with its business partners in Slice 1 and Slice 2. While for technical reasons and efficiency it decided to share the AMF function which manages the radio access and mobility.

<sup>1</sup> This in particular means that there is a secure tunnel end-point that would end in the network function and not in the slice.



## Host MNO

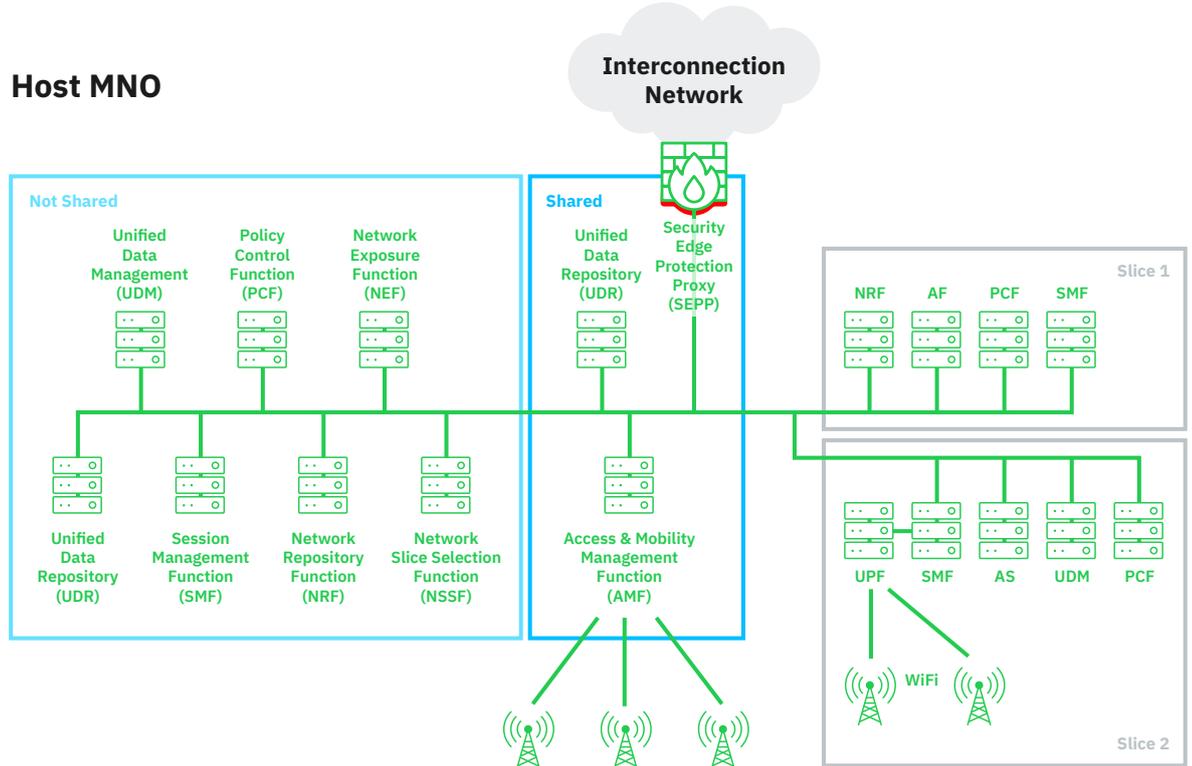


Figure 2: Example Slicing in the Core Network

While those Not Shared, Shared, Slice 1, Slice 2 “boxes” above indicate a logical / business-area type of separation, they are not completely separated on the signalling layer. All those network functions in the different “boxes” (i.e. the slices, shared and not-shared network parts) are connected to the SBA and its interfaces. This is because all those network functions need to exchange signalling messages with each other.

In some cases, two slices may want to communicate with each other, so they would use the common SBA. An example of such inter-slice communication is an automotive slice that wants to communicate with an entertainment slice for in-car entertainment purposes.

The network functions in the slice are managed by the MNO, but the content that is provided to the user and the details in the network function might be managed by the partner. A typical business case for such an architecture would be a sports stadium company that is managing the slice for their visitors and the content that is streamed to the user is coming from an advertisement company.

Network slicing allows the flexible ramping up of network functions, fast and dynamic service deployment, the logical separation of network functions and grouping of network functions. It also enables the support of different use cases and business models. For example, if a slice owner wants to run their own 5G user database or connect their local data serving network to the 5G network. This kind of targeted vertical support and flexibility is not possible in 4G.

### 2.2.1 Security Zones as a Consequence of Slicing

A slice is a logical group of network functions that communicate with each other to serve a specific purpose. From an architectural point of view, the MNO wants to group the network functions on the transport layer and the signalling layer into security zones, according to their trustworthiness or business arrangements.

One interesting point concerns partitioning - as a side-effect of TLS and IPsec usage, a form of de-facto partitioning comes into place, because only certain authenticated network functions can communicate with each other. While any sort of partitioning can be used - because it allows security zones to be created - this TLS/IPsec resulting partitioning does not provide a true security zoning of the network, especially when it comes to signalling traffic.



Nonetheless, the concept of security zoning is an important 5G security topic to introduce. We believe that to do true security zoning of the 5G core network, the following key protection points should be secured:

- between the network and the interconnection network,
- between network slices (inter-slice communication),
- between shared and non-shared network functions,
- between the dedicated network functions and the shared infrastructure, and
- between the 5G network functions and elements of legacy generations such as 2G, 3G and 4G.

How to protect those points best and what the open security issues are, will be outlined in chapter 3. First, we will explain slice properties and how current slice security is ensured when a user wants to utilize a service provided by a specific slice.

## 2.3 Slice Types and Properties

Slices are logical groups of network functions defined by a business purpose. The business purpose then defines the needed quality features of that slice like low latency, broadband etc. There are currently four formal slice types specified which we will introduce now.

### 2.3.1 Slice Types

The 3rd Generation Partnership Project (3GPP) currently defines in TS 23.501<sup>[2]</sup> the following four types of network slice types, based on their quality of service features:

1. massive Machine Type Communication (mMTC)
2. enhanced Mobile Broadband (eMBB)
3. Ultra-Reliable Low Latency Communications (URLLC)
4. Vehicle to X (V2X)

Each of the slices has its focus on a specific use case scenario:

- mMTC is intended to cover slices that service large amounts of Internet of Things (IoT) devices.
- eMBB intends to serve entertainment use cases. A typical example would be event streaming.
- URLLC can provide mission-critical networks or health applications with a suitable slice.
- V2X is focused on connected and self-driving cars.

While these use cases already cover a large range, we expect that more slice types will be defined in the future. One could envision that the URLLC slice type might be split into two - one for reliability, and one for low latency. A combination of eMBB features with low latency might also be envisioned e.g. for gaming applications, to form other new slice types.

Each slice instance is identified across the 5G core, the 5G RAN and in the User Equipment (UE)<sup>ii</sup> by a slice identity which is called the Single-Network Slice Selection Assistance Information (S-NSSAI) in TS 23.501. This identifier has two parts:

- Slice Service Type (SST) is a predefined value for eMBB or mMTC, etc.
- Slice Differentiator (SD) is an optional MNO specific value for differentiating between slices of the same type

Slice Service Type (SST)	Slice Differentiator (SD) - optional
1 (8 bits)	42 (24 bits)

Table 1: Example: S-NSSAI (Single Network Slice Selection Assistance Information)

<sup>[2]</sup> 3GPP TS 23.501, "System architecture for the 5G System (5GS)", version 16.6.0, 24. September 2020

<sup>ii</sup> Note: The User Equipment (UE) here refers to any mobile network using device, e.g. mobile phone, IoT devices etc.



An MNO may offer the same slice type to different verticals e.g. eMBB for different streaming providers or mMTC for different IoT service providers. The MNO can choose, if it wants, to populate the SD and what value to put there. Note, the MNO can also use non-standard S-NSSAI if it so wished, such as using their own non-3GPP defined SST value, or an MNO-specific and self-defined SST and SD.

Inside the core network, the S-NSSAI is used for traffic differentiation and QoS aspects, but also authorization, policy enforcement and potentially for routing. As a UE may belong to several slices, the concept of a group or list of slices was introduced, and this is (somewhat confusingly) called Network Slice Selection Assistance Information (NSSAI). There are different categories of NSSAIs. Typical categories used are Allowed NSSAI, Rejected NSSAI, Configured NSSAI, or a Requested NSSAI.

One key security takeaway is that there is no requirement, for an MNO, that each slice has a unique identifier or that the SD part is random. A random SD part would make guessing an S-NSSAI by an attacker much harder. The detailed implications of this will be discussed in section 3.

### 2.3.2 Slice Properties and Security Attributes

The GSMA Network Group (NG) specified the Generic Slice Template (GST) in their NG.116<sup>[3]</sup> specification. The GST is a set of characteristics for a type of slice or service. It is a generic set of mandatory and optional attributes. Those attributes are related to throughput, delay tolerance, radio spectrum etc. NG.116 v2.0<sup>[3]</sup> also contains the following security-related attributes, which focus mostly on physical and logical/virtual isolation. NG.116 divides isolation into the following 2 main types:

- Physical Isolation
  - Process and thread isolation
  - Physical memory isolation
  - Physical network isolation
- Logical Isolation
  - Virtual resource isolation – a network slice has access to a specific range of resources that do not overlap with other network slices (e.g. virtual machine isolation)
  - Network isolation - the network function is dedicated to the slice and that vertical customer, but virtual resources are shared
  - Tenant/Service Isolation – the vertical customer data are isolated from other verticals, but virtual resources and network functions are shared.

It is important to note that the NG.116 document has an end-to-end view i.e. it does not cover scenarios where part of the infrastructure is virtual and part of the network functions are physical nodes, or where some nodes are shared and some are dedicated to a slice. This means it does not provide different policies for different nodes, but only on an end-to-end basis for the whole slice. It remains to be seen if MNOs will always deploy a full end-to-end slicing scenario.

Other potential improvements are likely to be required in the future. For example, slice attributes related to security algorithm usage in RAN, as well as requirements for roaming security that are not part of NG.116 v2.0<sup>[3]</sup>.

### 2.3.3 Creating a Slice

Given all the security attributes and characteristics above, how will a slice with its security requirement come to life?

It starts with a vertical having its use cases and requirements. These are then translated into a Service Level Agreement (SLA). The MNO then analyses what network functions are needed and uses the corresponding

<sup>[3]</sup> GSMA NG.116, "Generic Network Slice Template", version 3.0, 22. May 2020



templates provided by GSMA NG.116 <sup>[3]</sup>, 3GPP TR 28.801 <sup>[4]</sup> and TS 28.531 <sup>[5]</sup>. The templates are filled in based on the SLA (including the slice specific isolation requirements) and then the filled-in template, called the NEtwork Slice Type (NEST), is used for the creation of the slice. The slice and underlying function are then tested, the network function on-boarded, it goes live, and then the hosting MNO enforces the QoS attributes (including security) for that slice (see figure 3).

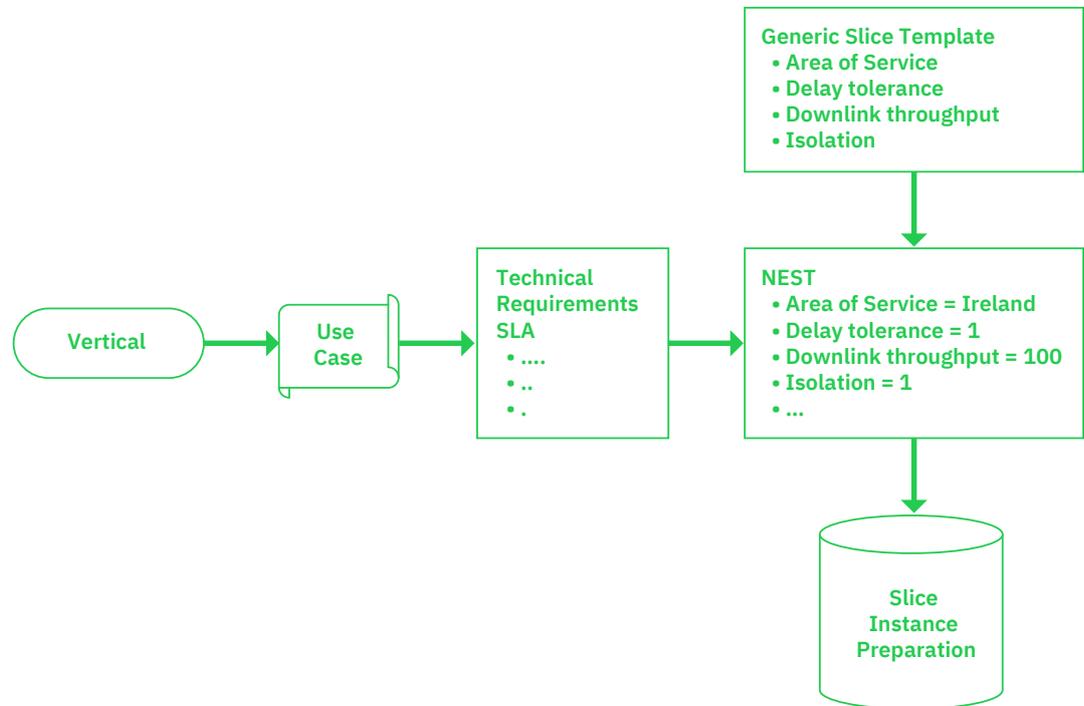


Figure 3: Path from Use Case to Slice Instance Preparation

Those are just the high-level actions - many smaller actions take place on the virtualization layer and in the Network Function Virtualization Management and Orchestration system. While virtualization can, if deployed, use TLS or IPSec for authentication between the network functions, those lower layers cannot take care of the authorization of service usage. TLS and IPSec do not provide fine-grained control on the signalling layer and sensitive information elements like UE location, data redirection addresses etc. The implications of this will be discussed in later sections.

<sup>[4]</sup> 3GPP TR 28.801, "Telecommunication management; Study on management and orchestration of network slicing for next generation network", version 15.1.0, 4, January 2018

<sup>[5]</sup> 3GPP TS 28.531, "Management and orchestration; Provisioning", version 16.7.0, 25, September 2020



## 2.4 Authorization in 5G Service Based Architecture

In the 5G SBA, every network function can communicate with every other network function, which is different from 4G, where each interface had to be “switched on” individually. For this reason, the authorization function, called NRF, was introduced. It helps network function consumers to discover services, authorizes requests and issues “access tokens” to services.

The new architecture can also potentially create some load-handling scenarios that need to be dealt with - how does the network deal with misconfigurations and how does it co-ordinate the overall availability of services in the systems? Due to this, 3GPP has recently introduced in Release 16 the concept of the Service Communication Proxy (SCP). That is a new network function in the SBA architecture, and it can assist the handling of inter-NF messages within the network. We will describe now how both the NRF and SCP handle authorization for slicing.

### 2.4.1 Network Repository Function (NRF)

As each network function in the SBA can contact every other network function, the questions arise:

- How can one network function discover the services offered by another network function? and
- How does the network function receiver of a message from another network function, determine if the request for a service is authorized?

The answer to these questions lies in the NRF. The NRF is defined by 3GPP in TS 29.510 <sup>[7]</sup> and it offers three main services:

1. Network Function Management, which allows network function instances in the same network to register, update and de-register their profile in the NRF.
2. Network Function Discovery, which allows a network function to discover the services of other network function instances in the same network. It also supports one NRF querying an NRF in another network on behalf of a network function.
3. OAuth2 Authorization service, where the NRF issues a token to the requesting network function, which the requesting network function can use to prove that it is authorized to consume the service of another network function.

For those services, the NRF maintains a database of all network function instances, their status and the services offered by them in the local network. Each network function has a description called an NFProfile (TS 29.510 <sup>[7]</sup>). This profile for a network function may contain many items related to slices. Important elements of the **NFProfile** are:

- Supported NSSAI: A list of S-NSSAIs (slice identities) supported by the network function, which also contains the slice types supported by this network function. Note supported is not equal to allowed.
- Roaming NSSAI: To cover the roaming case, the NFProfile may contain a list of S-NSSAIs supported by the network function for the other 5G networks (in case they are different from the previous bullet).
- Allowed NSSAI: A list of S-NSSAI (allowed S-NSSAI). This list is used for access control and contains the identities of slices that are allowed (not just supported) to access the services of a network function.

From a security perspective, it is also interesting to know that the NFProfile information can also contain dynamic load information TS 29.510 <sup>[7]</sup>. This information is stored in an Information Element (IE). The NRF also provides a notification service that informs a network function that has subscribed to a particular notification service, if an event occurs like the UE is reachable, changes in location of a UE etc. The notification service can also indicate an overload situation in the home MNO, called Public Land Mobile Networks (PLMNs). The notified network function in the other PLMN can then take measures to avoid further overloading the network function in the home PLMN.

<sup>[7]</sup> 3GPP TS 29.510, “5G System; Network function repository services; Stage 3”, version 16.5.0, 25. September 2020



It is required in TS 33.501<sup>[14]</sup>, that the network function and the NRF mutually authenticate each other using TLS or IPsec on the lower layer of the Open Systems Interconnection (OSI) stack model, and that the NRF may provide authentication and authorization to network functions for establishing secure communication between each other. But lower layer security only provides communication security between two network functions and does not provide authorization of service access or authorization to access individual sensitive information elements. Once a network function is regarded as trusted to use a service, it is trusted with everything that this service may encompass. The diagram below is an example of how a service consumer e.g. a network function that needs to know the location of a UE, can register itself with the NRF and then request an access token to obtain the location information from an AMF.

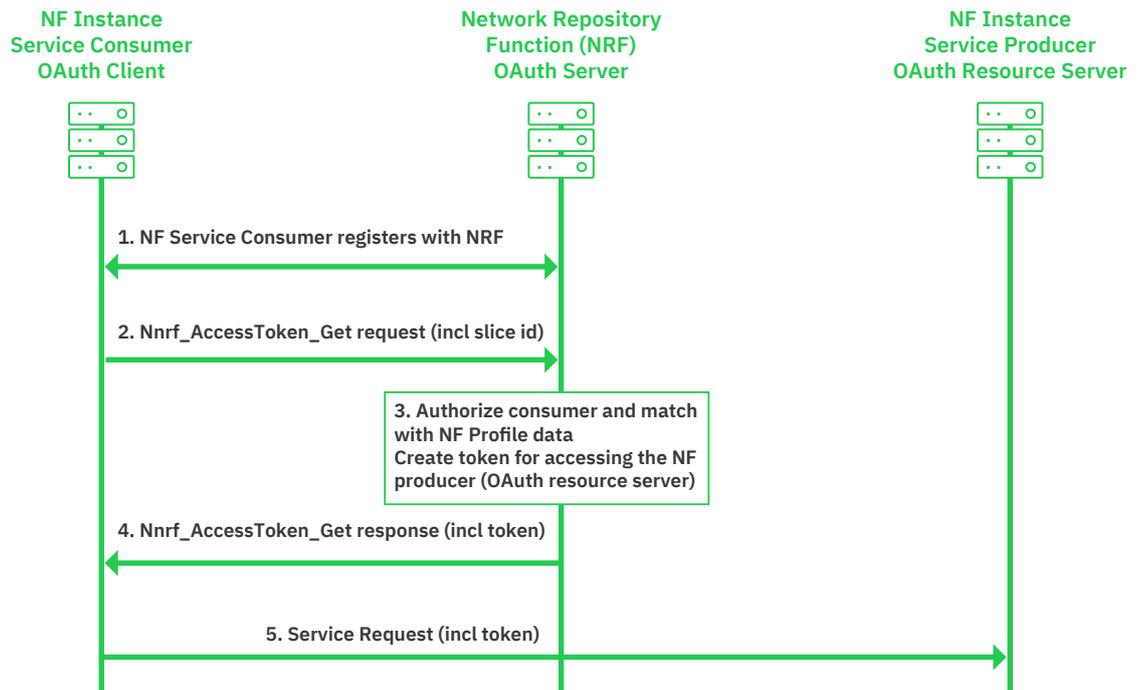


Figure 4: Network Function Consumer Obtaining Access Token from NRF for Service Usage

The NRF acts as an OAuth Server and the requesting network function as an OAuth Client, the network function which offers the service (aka the network function producer) is the OAuth resource server. In figure 4 above, the network function on the left, which wants to consume a service, registers first with the NRF in the centre. In step 2, the NRF receives from the NF consumer the request for a token for the service producer. The NRF then validates in the third step if this NF function instance, which wants to consume a service for a specific slice is allowed to do so based on the slice specific information in the NFProfile. This check uses the slice identity (S-NSSAI) provided by the NF consumer and matches it against that stored in the NFProfile of the NF service producer. If this check-in step 3 is successful, the resultant issued token to the requesting network function access services is issued in step 4. The NF consumer then contacts the NF producer on the right in step 5 to obtain the service via the service API and includes the token it received from the NRF.

The NRF authorization mechanism described above allows slice specific service authorization - it is important to note it is not information element level specific authorization e.g. load or user-identity. The authorization mechanism relies on the slice identity information presented by the network function consumer when the network function contacts the NRF to obtain a token to access the network function resources. But a rogue network function may lie about this and present a slice identity of another slice. The specification does not mention any cross-checks of whether the NF consumer presented slice identity belongs to this NF consumer. The presented slice identity is only used to identify the correct NF resource (service producer). Details of how this vulnerability can be exploited can be found in section 3.1.3, which explains how to access the resources

<sup>[14]</sup> 3GPP TS 33.501, "Security architecture and procedures for 5G System", version 16.5.0, 16. December 2020



of another slice and how to perform a DoS attack on another slice. In section 3.1.5 we explain how to extract user specific information like location from another slice.

Note, as explained earlier, there is an NFProfile that is used by the NRF to determine supported, roaming and allowed slices for the NF that offers the resource. But it is not used for validating the correctness of the requested (i.e. of the presented slice identity by the network function consumer).

From a slicing perspective, each slice could have its own NRF for authorization of service access, in their “box”, to network functions residing in its slice. This may result in several NRFs in the same network and the need to transfer requests from one NRF to the next one until the right one is found. As an example, below in figure 5 the first NRF1 in a part of the network that is not shared with slices, forwards or redirects to the next NRF2 function, which could be shared between slices, until it reaches NRF3 that is slice specific.

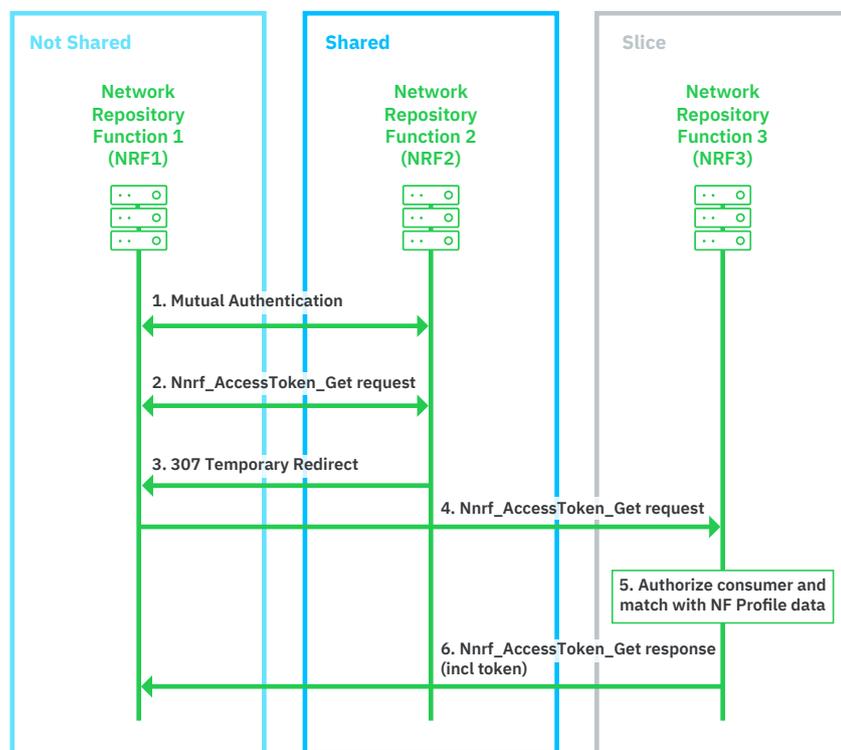


Figure 5: Slice Specific NRF - Forwarding of Requests from TS 29.510 [7]



As authorization to access services is an important security control to avoid attacks coming from the Interconnection Networks (IPX), we will now look into how authorization works across PLMN boundaries, i.e. between 5G networks. In the roaming scenario shown in figure 6, the network function consumer instance in the visited network that wants to use a service in the home network contacts the NRF in the visited network (vNRF) and registers with the vNRF in step1. The network function consumer would not be aware of the NRF in the home network (hNRF).

The visited network vNRF and the home network hNRF authenticate each other and communicate (step 2) via the SEPP of the visited network and the SEPP of the home network. In step 3 the NF consumer would request from the vNRF the access token. In step 4 the vNRF would authenticate the NF consumer (client). In step 5 the vNRF requests on behalf of “its” NF consumer the token from the hNRF. In step 6 the home network hNRF would authorize the client and issue a token which is then returned in step 7 to the vNRF. The vNRF then forwards this token to the consumer. Below are the individual steps that take place, when a network function in the visited network wants to consume a service in the UE’s home network.

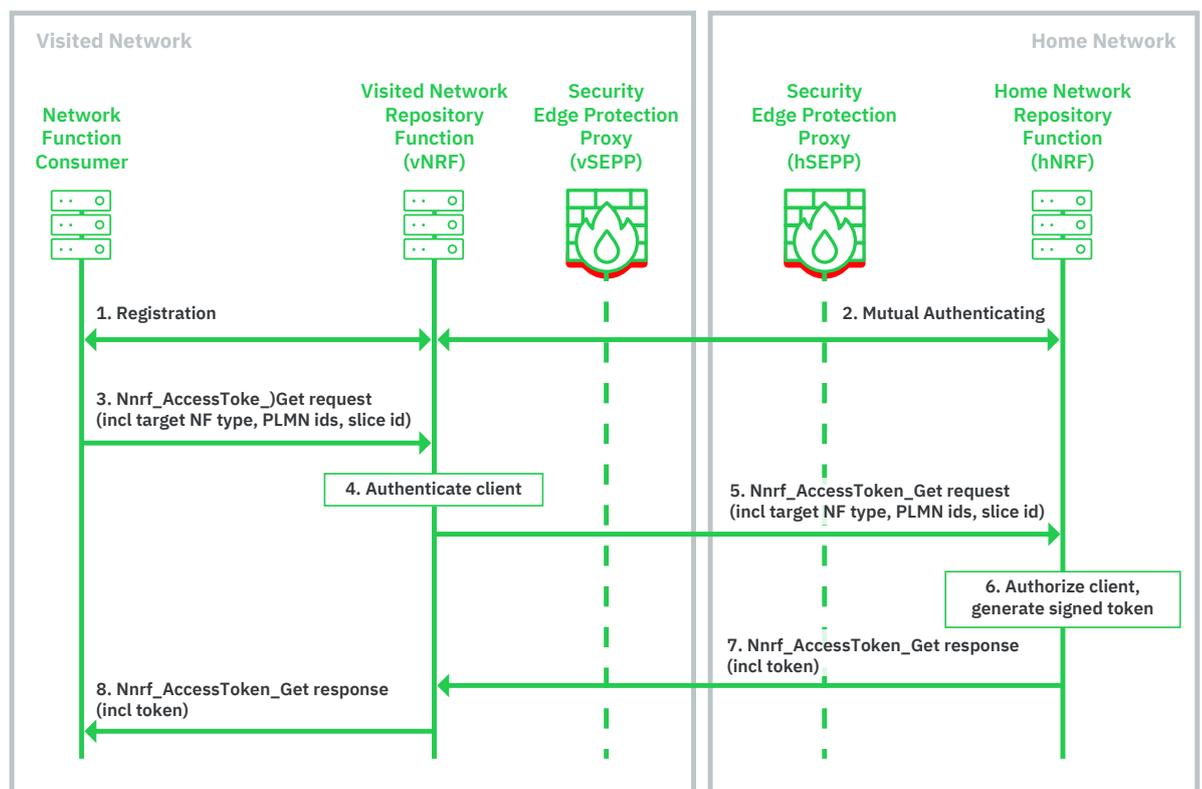


Figure 6: Authorization Token Request in Roaming Case

Note that authorization across PLMNs can also include the load information element in the NFProfile description and the Overload Control Indicator (OCI) in the HTTP header (TS 29.500<sup>[8]</sup>). The OCI allows one network function to indicate overload situations to another, even across networks boundaries. Such an indication can be put into a service request or other types of messages. The receiving network function is then aware that the sender is suffering under load and can take measures to reduce the load. Such a load indication needs to be secured properly to avoid Denial of Service (DoS) attacks. For details of how this vulnerability can be exploited, see section 3.1.3.

To deal with the complexity and dynamic nature of the SBA, with NFs communicating on it, the 3GPP Release 16 introduced a new optional element called the SCP. This interacts closely with the NRF and can act on behalf of the network functions. We will outline the details of the SCP and how it supports slicing security next.

<sup>[8]</sup> 3GPP TS 29.500, “5G System; Technical Realization of Service Based Architecture; Stage 3”, version 16.5.0, 24. September 2020



### 2.4.2 Service Communication Proxy (SCP)

The SCP is a new type of element in the 5G architecture. As stated, it was recently introduced in Release 16 of the 3GPP architecture, in TS 23.501 [2] and TS 33.501 [14], as an optional enhancement to make the management of the SBA easier. Like the SEPP, it sits on the SBA bus (see figure 7). The SCP's focus is on network internal communication and, in general, to ease practical deployments so that not every NF has to perform every task itself. It stores a general overview of the network topology and can perform tasks like:

- Indirect service communication between different network functions
- Delegated discovery to discover, on behalf of a network function, the requested network function which could provide the requested service
- Message forwarding and routing to destination network function and control actions
- Message forwarding and routing to a next-hop SCP, in case there are several, or for roaming
- Communication security (e.g. authorization of the NF Service Consumer to access the NF Service Producer API)
- Network reliability protection i.e. load balancing, monitoring, overload control, etc.

It should be noted that an SCP must support at least one of these features and may support several of them. The main point of the introduction of the SCP was to have an architectural element that can support and steer the communication within the SBA.

### Service Based Architecture (SBA) for indirect communication

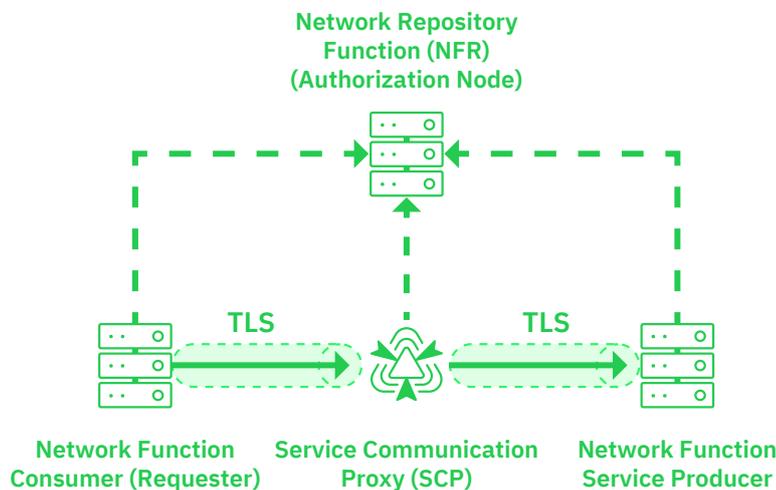


Figure 7: Service Communication Proxy

The SCP might also interact with network functions to resolve groups and addresses. Some of the interactions like the NRF procedures for authorization of a slice are specified within the 3GPP standards. Others like monitoring, overload control, load balancing, are not specified and left to the implementation. For example, the traffic steering interaction is not specified by 3GPP, as it is not needed for interoperability, but technically it could be based on the slice identity.

The SCP sits in the “middle” of the SBA and with specific enhancements can act potentially as a guard for network internal traffic and security zones, slice isolation and security filtering. We will describe in chapter 3 how the SCP functionality can be enhanced to improve the security of the whole system.



## 2.5 User Equipment Access to a Slice

The 3GPP security has a strong focus on the attack scenario that only authorized user equipment (UE) can access and use the services provided by a certain network slice. For that purpose, 3GPP TS 29.531<sup>[16]</sup> defines a Network Slice Selection Function (NSSF), which contains the NSSAI. This NSSAI is essentially a list of slice identities i.e. it is a list of S-NSSAIs. A common vertical use case is that a UE would have access to a specific data network, such as a private network or factory network. The subscription information of the UEs for each slice would also contain different Data Network Names (DNNs) which identify the data networks the UEs are allowed to access and which would belong to “their” slices.

The DNNs are the 5G equivalent of Access Point Names (APN) for 4G, and for vertical deployments, these could be the factory floor network or a specific private network. A UE is pre-configured or provisioned with a default NSSAI, which if you recall is a group of S-NSSAIs (slice identities) of its home network.

A UE can use services of several slices for example a private network or the Internet. If the UE is roaming, the home network may have the NSSAI for the visited network already configured in it. If there is no such list, then the NSSAI of the home network is used. If the UE has the NSSAI for the visited network configured - that the visited network does not recognize - then the visited network can update the allowed NSSAI slices with a mapping of its own corresponding S-NSSAIs (slice identities) e.g. based on SST. The interactions around slice specific roaming security have been recently defined by GSMA in NG.113<sup>[15]</sup>.

The S-NSSAI slice identity is the key cornerstone for authenticating and authorizing the UE access to a slice. There are two types of mechanism for controlling UE access to a slice:

- Simple slice access done during the registration of the UE
- Slice specific access which requires an extra authentication step (e.g. for Intranet access of private networks)

The first relies only on “normal” network authentication, and the slice identities (S-NSSAI), and it is also the method used when a UE is roaming.

The second is intended to use additional authentication using the Extensible Authentication Protocol (EAP). A typical use case for the extra authentication mechanisms would be that a company wants to have an additional authentication step before a UE can access the factory network and therefore would run an authentication server for the EAP authentication. Both are covered in more detail below.

### 2.5.1 Simple Slice Access

The access to a slice is part of the normal registration procedure in 5G and defined in 3GPP TS 23.502<sup>[13]</sup>. The following steps are described in figure 8:

1. The UE sends to the RAN in the UE registration request a list of S-NSSAI (requested NSSAI) and potentially a mapping of requested NSSAI, which assists, in case of roaming, to find the correct slice for the UE. The RAN network is not aware of the subscription data for this UE.
2. The RAN performs an initial AMF selection based on the information provided by the UE. This selection can be based on a potential AMF address provided by the UE or based on Radio Access Technology (RAT) and requested NSSAI. The RAN may also have a local configuration if insufficient information is provided or the data provided is invalid.
3. The RAN sends the registration request to the initial AMF. This message also contains the requested NSSAI and the mapping if it was provided.
4. The initial AMF must validate if the user is allowed to access those S-NSSAI. For that, it contacts the UDM to request UE’s Slice Selection Subscription data.
5. The initial UDM (after potentially having fetched the UE’s Slice Selection Subscription data from the UDR) provides the requested data to the initial AMF.

<sup>[13]</sup> 3GPP TS 23.502, “Procedures for the 5G System (5GS)”, version 16.6.0, 24. September 2020

<sup>[15]</sup> GSMA NG.113, “5G Roaming Guidelines”, version 2.0, 28. May 2020

<sup>[16]</sup> 3GPP TS 29.531, “5G System; Network Slice Selection Services; Stage 3”, version 16.4.0, 25. September 2020



6. The AMF now has the data to cross-check if the UE is allowed to access the slices he requested. The AMF now knows which slices the UE is subscribed to from the data provided by the UDM.
7. The initial AMF might not be able to serve all the S-NSSAIs from the request NSSAI that the UE is subscribed to. In that case, it sends a network slice selection request to the NSSF. It contains among other parameters the requested NSSAI, mapping, subscribed S-NSSAI. The NSSF validates the request according to TS 23.501 <sup>[2]</sup> 5.15.5.2.1 option B. based on the provided information, tracking area, potential roaming scenario, and configuration. The NSSF may now need to contact the NRF to discover the target AMFs for this UE.
8. The NSSF contacts the NRF to request a list of candidate AMF(s) and includes the list of S-NSSAI it deems suitable according to the procedure in TS 23.501 <sup>[2]</sup> that the candidate AMF needs to support.
9. NRF discovers the suitable AMF instances and returns a candidate list of AMF(s) to the NSSF.
10. The NSSF returns to initial AMF the allowed NSSAI, optionally the mapping of allowed NSSAI and the target AMF set or, based on configuration, the list of candidate AMF(s). There can be more NSSAI lists if there is a second access type and slice specific NRF information for the case we described in Figure 5.
11. The initial AMF received the candidate AMF. For the case that it does not have the candidate AMF instance address stored it needs to contact the NRF for discovery. The initial AMF now has two possibilities: either it can redirect the UE to the new target AMF, or it can inform the target AMF to take care of the UE. This decision is based on local configuration and subscription information. In Figure 8, the initial AMF decides to reroute the message via the RAN. For this, it sends a Reroute NAS message to the RAN which includes the information needed for the new target AMF, the full Registration Request.
12. The RAN sends the Initial UE message to the new target AMF and indicates reroute due to the slicing information provided from the NSSF via the initial AMF in the previous message.
13. The new AMF that serves the UE requested slice now continues with the normal Registration procedure according to TS 23.502 <sup>[13]</sup>.

If an SCP is deployed it may assist in any of the discovery procedures. It should be noted that TS 23.501 <sup>[2]</sup> and TS 23.502 <sup>[13]</sup> have many specific cases for the slice selection and the usage of the NRF. We describe only one scenario here; the key security take-away is that 3GPP standardization protects well against the attack scenario where a UE is malicious and tries to gain unauthorized access to a slice. From a high level, it can be seen that the control complexity and granularity for the UE access is much higher compared to the network function to network function control approach.

The target AMF also uses the slice identity to determine the correct SMF to set-up the data sessions for the UE. The data session (PDU Session Context) in the SMF, AMF and the RAN contains one and only one slice identity (S-NSSAI), which means that the data session is bound to a particular slice. The policy function PCF can inform the AMF if the list of allowed or configured slices for the UE has changed.

The overall basic concept is that the UE presents the slices it is configured to use and then the network cross-checks with the subscription database during the registration if this is correct and then triggers the needed network functions for setting up the access. It is worth noting that after the AMF has made the UE / slice identity check, no further cross-checks like that take place by the NRF or other network functions.

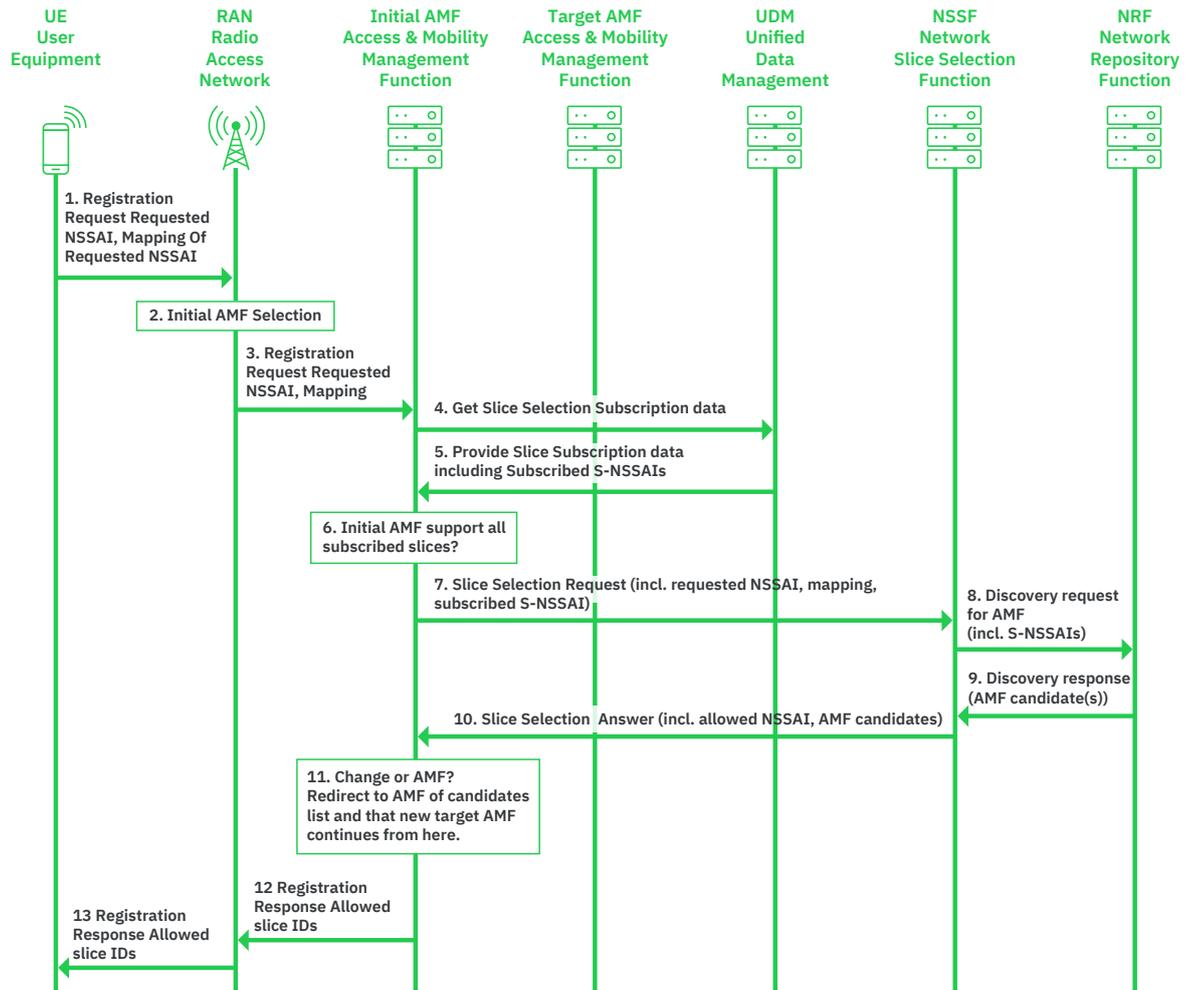


Figure 8: UE Slice Access Security

### 2.5.2 Slice Specific Authentication

In the case that the slice use case requires an additional level of security for the UE to access the slice, there is the possibility that an additional authentication of the UE may be required. This requirement is part of the subscription information of the UE.

During the registration procedure of the UE, the slice identity S-NSSAI is used to trigger the usage of the EAP framework (as defined in TS 33.501<sup>[14]</sup>) with an AAA Server (AAA-S) which can be hosted by the MNO or by a third party, such as a company's AAA-Server for intranet access<sup>iii</sup>. The UE needs to support the Network Slice Specific Authentication & Authorization (NSSAA) feature in the UE to be able to perform slice specific authentication with the AAA server, which was only introduced in 3GPP Release 16.

When the UE is connecting to the network and the registration procedure is triggered in the AMF (step 3 Figure 8 above), then the AMF decides that the network slice-specific authentication and authorization is required for that slice based on the S-NSSAI slice identity. The AMF then takes the role of the EAP Authenticator when communicating with the AAA Server. The NSSAAF (Network Slice Specific Authentication & Authorization Function) performs the interworking of any AAA protocol and the protocol supported by the AAA Server. The NSSAAF is an interworking function to plug in external authentication servers to the SBA of the MNO. Here we again see a server (i.e. the AAA Server for additional authentication) that is connected to the core network, which could be a third-party node.

<sup>iii</sup> This is defined in TS 23.502 [13] and only recently with the closing of 3GPP Release 16 final details were added.



As a summary of section 2, the core network relies on a list of S-NSSAI slice identities concerning the UE identity, which is then matched against policies stored in the UDR or UDM. This information is used to prevent unauthorized access from the UE to a slice. But one thing that becomes clear from this is there is no means to validate if a network function is presenting the correct S-NSSAI slice identity, as the network functions assume that the check between UE identity and S-NSSAI slice identity was performed when the UE was connected. These kinds of missing checks can lead to the security challenges we discuss in the next chapter.

### 3. Security Challenges for 5G Slicing and Countermeasures

In this chapter, we will discuss the security challenges from slicing and bringing external parties into the network, as well as covering aspects like legacy interworking and configuration risks introduced through the increase in complexity. We will describe how an attacker may exploit missing checks or careless configurations to gain unauthorized access to resources. As a potential countermeasure approach, we discuss how many of the new security features can also be extended to divide the network into well-secured security “zones”.

#### 3.1 Security Challenges

The 5G SBA offers many security features which includes lessons learned from previous generations of network technologies. But on the other hand, 5G SBA is a completely new network concept that opens the network up to new partners and services. These all lead to new security challenges.

##### 3.1.1 Security Challenges Through Legacy Interworking

Mobile networks evolve, and each evolution is more complex than the previous generation. This is also true for 5G. Up to the time of writing (Mar 2021), there has been multiple news stories of 5G rollouts, however, in truth, many of them are only partial 5G radio network deployments<sup>[17]</sup>, and not the rollout of 5G core networks. Even when MNOs move to a 5G core, it is prudent to assume that they will continue to support many legacy, pre-5G functions in their own network. This can be due to migration plans or due to specifications with legacy integration e.g. for SMS over IP, SMS over Non-Access Stratum (NAS). Also, not every MNO worldwide will move at the same speed - MNOs will have to support interworking with roaming partners that do not yet have a 5G core and that supports 5G roaming.

<sup>[17]</sup> Fierce Wireless, “‘Real’ 5G relies on 5G NR, Standalone architecture: Special Report”, 24. May 2020, <https://www.fiercewireless.com/wireless/real-5g-relies-5g-nr-standalone-architecture-special-report>

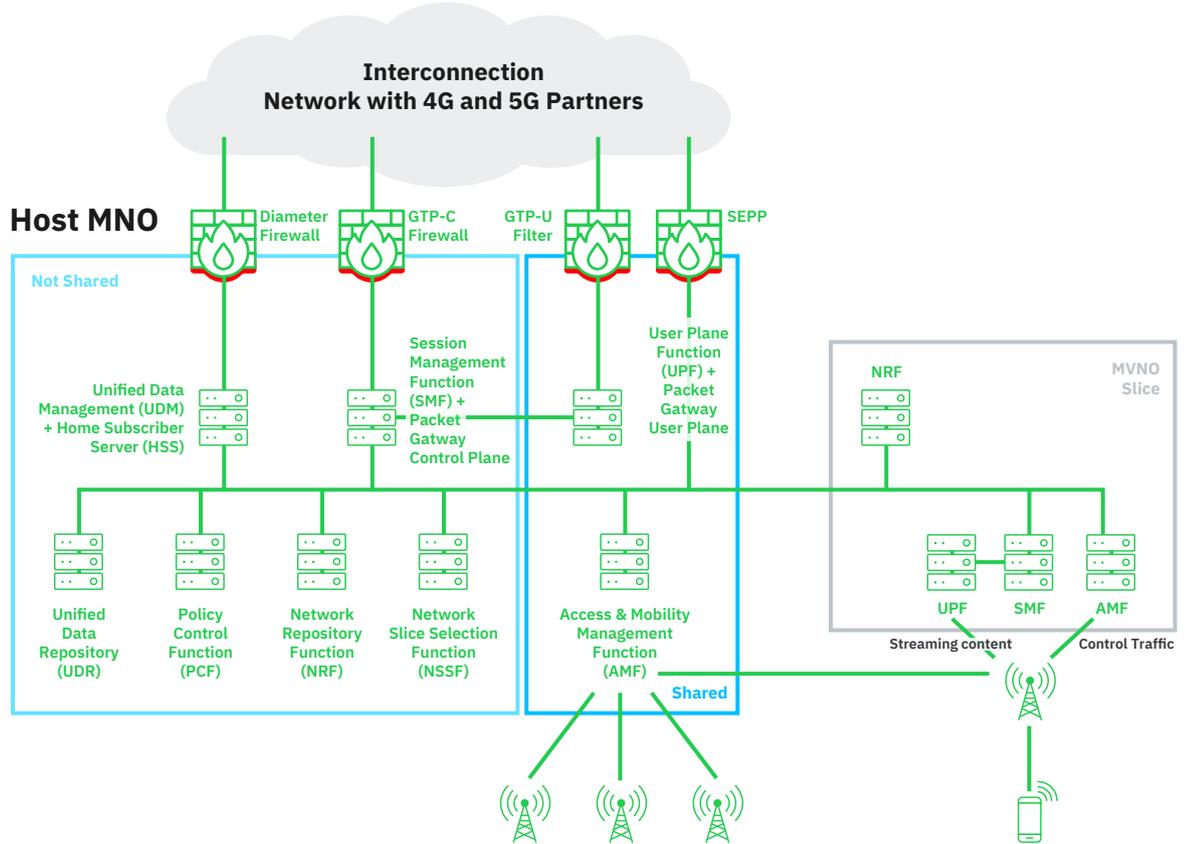


Figure 9: Legacy Roaming and Interface Protection

Migration of networks to a new generation requires careful security considerations and are illustrated in Figure 9. In previous AdaptiveMobile Security research on this subject, we showed <sup>[12]</sup> that mobile core network attackers use whatever protocol and network generation that gives them the result they desire. As a result, we outlined in our white paper how to secure the migration path from legacy protocols towards 5G <sup>[6]</sup>.

In a mixed architecture where some elements are 4G and some 5G, interworking functions will enable communications across generations. As 4G does not have a slicing concept, how will slicing security be dealt with e.g. slice specific service authorization, when the 4G interworking function converts communication from that 5G slice to the 4G network? And vice versa there are also questions. The 5G interworking function would potentially interface the whole 4G legacy node as one service to 5G network functions. Would the 5G slice be allowed to contact legacy nodes in the 4G network? These are issues which have not yet been addressed in the standards as they are specific to the migration approach an MNO takes.

### 3.1.2 Security Challenges Through Increased Complexity

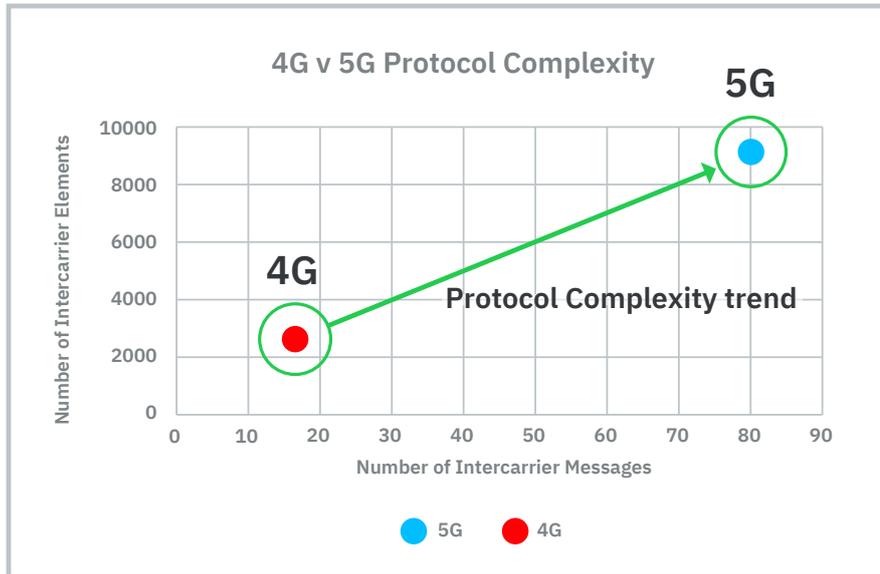
Networks provide more and more features and services but configuring them is not an easy task and mistakes will be made. Already today we have seen core network nodes that have their control traffic port open to the internet e.g. in the past on shodan.io it was possible to find GGSNs with GTP-C ports open. Also, the progress from 4G to 5G represents a massive increase in protocol complexity. To show this visually, as part of this white paper, we analyzed the protocol structure of 4G and 5G core network protocols, as a simple method to compare complexity. Figure 10 shows that:

- there are 4.7 times as many types of commands (messages) that can be sent inter-MNO over 5G, compared to 4G
- Over 3.4 times as many information elements (attributes) sent inter-MNO over 5G, compared to 4G

<sup>[12]</sup> Adaptive Mobile Security, "SIMjacker White Paper", <https://simjacker.com/>



This shows a very clear trend in complexity between 4G and 5G. By these metrics, 5G is several times more complex. This has obvious security implications because every single one of these commands needs to be inspected, and many elements approved, to prevent illegal or unwanted activity to be sent to and from a network. This becomes more difficult the more commands and information element types are received.



*AdaptiveMobile Research: 4G v 5G Complexity*

Figure 10: Protocol Complexity Trend

The above is based on inter-MNO message flows. Also, in 5G there are over 400 types of commands and nearly 50k different information elements that can be sent intra-MNO. This means that the amount of information elements that can be sent inside one MNO network is even larger.

As well as this, complexity is continuing to increase. The above are taken from 3GPP Rel-15, for Rel-16 there is an even higher number of command and information elements that are sent both intra and inter-MNO.

Not only has the complexity increased, but also the responsibility of the MNO has increased concerning his or her slicing customers. As the hosting MNO is responsible for the slice management, they must protect their slice customers and not just their own network against attacks. Below, in figure 11, is a graph of the growth of MNO and MVNOs over the last 15 years, and a 5 year projection from now. However, if we make a conservative estimate that in 5 years each MNO will have 5G and will support at least 5 slices, then this implies that there will be many more entities accessing the interconnection network over the next few years. These will all require proper configuration and filtering, which has multiple implications.

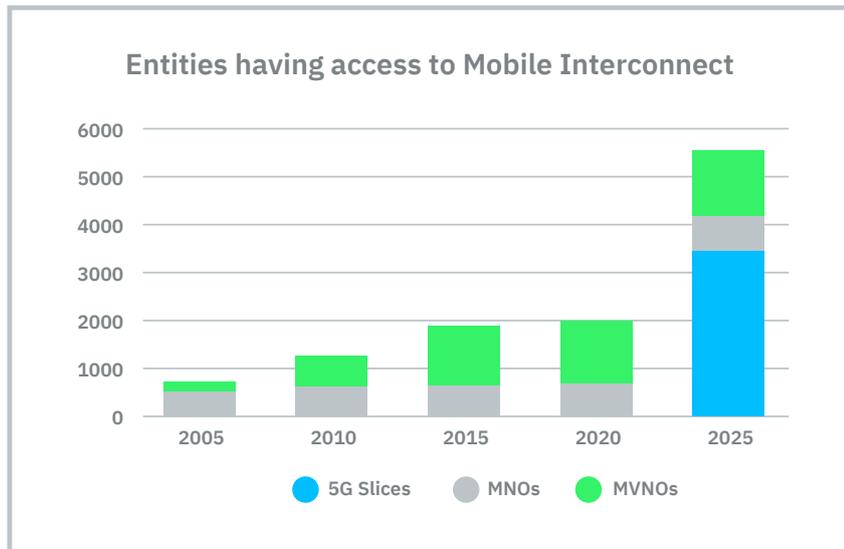


Figure 11: Conservative Estimate of Complexity Increase for Interconnection Security

The existence of several new entities - slices - within an MNOs' network can cause several security model implications. In the past MNOs were probably less concerned about activity leaving their network, as they would assume that it was "good", and traffic received from others is "suspicious". However, as covered in our previous research, with a set of new slice instances within a network, an MNO may need to monitor much more closely whether rogue slices may exist within their network, as they could send traffic outwards that would damage the host MNOs reputation, as well as directly affecting other MNOs.

Also, while "five times more entities" does not sound huge, it is a substantially higher load for the operational teams inside an MNO that perform the configuration and operational tasks. This can easily lead to oversights in configuration or "shortcuts" in security, which brings us to the next challenge.

### 3.1.3 Security Challenges Through Configuration Mistakes and Missing Layer Matching

With all the new network functions and services in 5G, roaming and legacy interaction will become quite complex. The configuration will be key, but assuming that MNOs will always configure everything correctly is dangerous, especially as knowing what is 'correct' may not be possible. One aspect highlighted in this section is slicing and configuration, and the ability for slices to do more than they should be able to do.

#### 3.1.3.1 Vulnerability: Malicious Access to Different Slice by Modifying Slice Differentiator

The concept of slicing is driven by the business case. One key aspect of the S-NSSAI slice identity is that it is used to provide the right "features" to the UE so that the service delivery and UE user experience is as expected. For these reasons, the S-NSSAI slice identity contains the mandatory SST (Slice Service Type) which is a predefined value e.g. for eMBB or mMTC, depending on the use case. A second optional part exists in the S-NSSAI - the Slice Differentiator (SD). These are explained in more detail in section 2.3.1

When an MNO starts deploying slicing, will the SD value be set every time, e.g. for the first slice? If no value is set for the first slice, will somebody remember to set an SD value, when the second slice of the same type comes? Will it be random or something obvious like a company name, where an attacker can guess the differentiator based on a press release? As there are currently no requirements to make the slice differentiator mandatory and random, guessing or "brute forcing" is a potential risk. If the differentiator is not set or is guessable, then it opens the possibility that a rogue network function or rogue slice, e.g. from a compromised partner, can use this information to gain unauthorized information or resource access.

We look at how this attack works in figure 12. The NRF specification TS 29.510<sup>[7]</sup> describes that the NF and the NRF need to mutually authenticate each other. We assume that a network function or the slice has gone rogue.

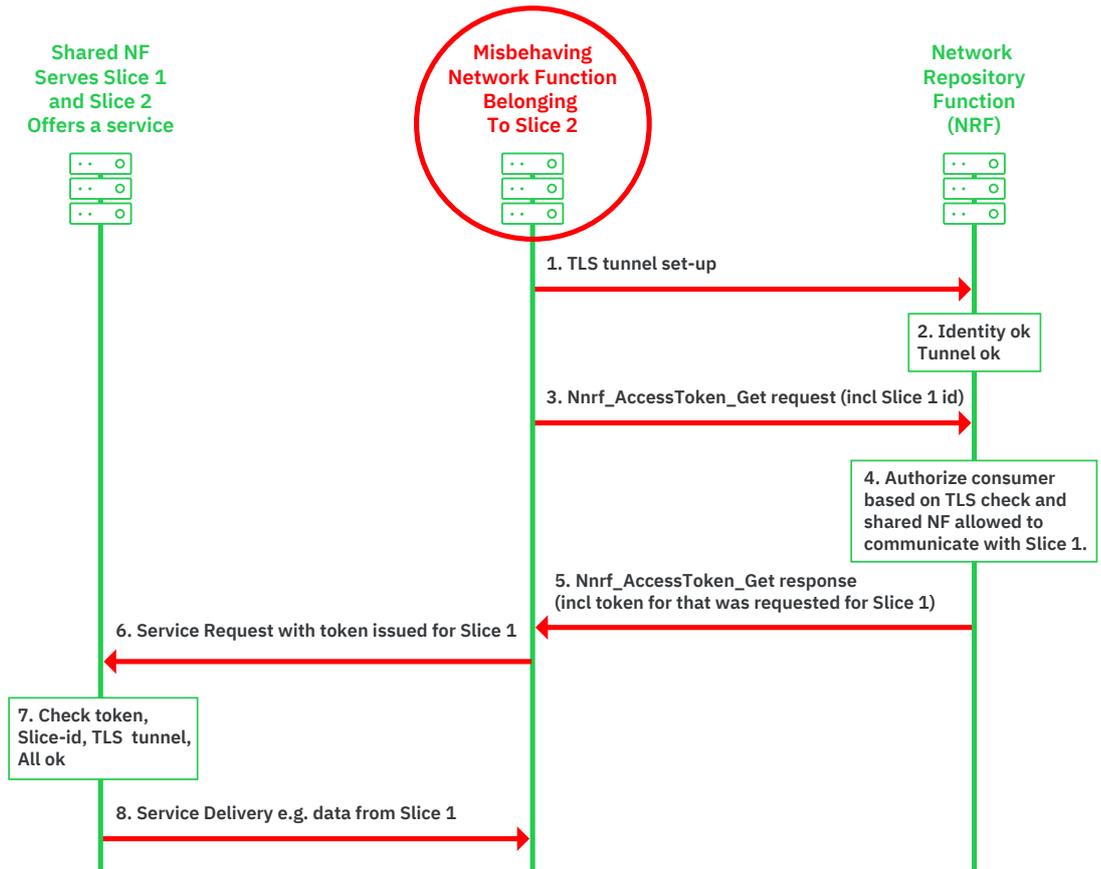


Figure 12: Rogue Slice Gains Potential Access to Other Slice Resources

In our example, of a simulated attack, the rogue network function belonging to Slice 2 would first establish a TLS connection to the NRF (step 1). The rogue slice 2 is allowed to contact the NRF as it belongs to Slice 2 (step 2). Then the rogue network function would request in step 3 from the NRF a token to access Slice 1 on the shared NF server. This request would contain a slice identity “Slice 1”.

In step 4 the NRF would check<sup>iv</sup> if the network function that the rogue slice wants to access is allowed to be accessed [7], but as the target network function is a shared network function for Slice 1 and Slice 2, the performed check would not result in dropping the request. The rogue slice is allowed to access the shared network function and a valid token for Slice 1 and the shared NF would be generated. The token contains the AuthenticationTokenClaims, which define the scope of the token. While there is the possibility to have *additional scope information* in the latest version of TS 33.501 [14], section 13.4.1.0, it does not give any specific formatting and details on how the requestor slice can be represented in this *additional scope information*. This implies, that there is currently no interoperability between vendors for the *additional scope information*. We assume that the *additional scope information* is used for conveying slice specific authorization information, even if it is currently not fully formally specified by the 3GPP.

This access token would be sent to our rogue Slice 2 by the NRF in step 5. In step 6 the rogue network slice would present to the shared network function the token for Slice 1 and as the shared network function would rely on the NRF to have performed all necessary authorization checks, would accept the token (step 7). If in step 7, transport/IP layer security is used between the network function belonging to Slice 2 and the shared network function, this would be a valid connection (see definition of shared network function). As a result, in step 8 the service would be delivered, which could be personal data like the UE location.

<sup>iv</sup> Note, there does exist a NFProfile, however that profile is only used for discovery of a service and is not used in TS 29.510 [7] for issuing the access token. The NFProfile is only part of the discovery of the consumer i.e. Nnrf\_NFDiscovery API and not of the Nnrf\_TokenRequest API. Even, if it would be used, it would not be used in verifying that the resource requestor and slice identity really match.



The underlying problem is that no layer matching is mandated by the specifications. As there is no matching between layers, the NRF would only see, on the lower transport/IP layer, an “authenticated partner” and, on the upper signalling layer, a valid slice identity and service request. There is no cross-check that the slice identity in the request matches the slice identity used for the TLS tunnel. As a result, the NRF would issue an authorization ticket to the rogue Slice 2 to use services on Slice 1.

There are also potential security concerns in the roaming scenario, as there seems to be nothing stopping a rogue slice in one MNO obtaining a token for service for his slice in another MNO.

### **3.1.3.2 Vulnerability: Specifying Different Slice in HTTP Header for DoS**

Another challenge relates to the HTTP header of 5G messages sent between network functions TS 29.500<sup>[8]</sup> and the validation of the elements within the messages. 3GPP has an overload control indicator header information, which is part of the HTTP header and can be used to indicate overload from one network function to another during normal service operations.

Assuming in figure 13 that a rogue Slice 2 wants to run a DoS Attack against Slice 1. Step 1 – A network function in Slice 2 authenticates correctly. It requests in step 2 an access token from the NRF for a service and slice 2. Then in step 3 the rogue network function sends an HTTP based service request containing the valid token to the service providing shared network function. Step 1 to step 3 are valid and normal messages for a service usage provided by the shared network function for Slice 2.

However, in step 4, the rogue network function can then create an HTTP message with a service request for Slice 2 that includes in the overload indicating 3gpp-Sbi-Oci header information the slice identity (S-NSSAI) of Slice 1 (TS 29.510<sup>[7]</sup>). The receiving shared network function would validate the token for the usage of the API, which is issued for Slice 2, which would be correct. The overload header information 3gpp-Sbi-Oci would potentially not be further cross-checked against the token used for the API service access. Therefore, in step 5, the service providing network function would assume that Slice 1 was overloaded and should not be contacted, for example, with notifications. This overload situation would be stored with the corresponding time information given in the header. The service would be delivered to Slice 2 as requested in step 6.

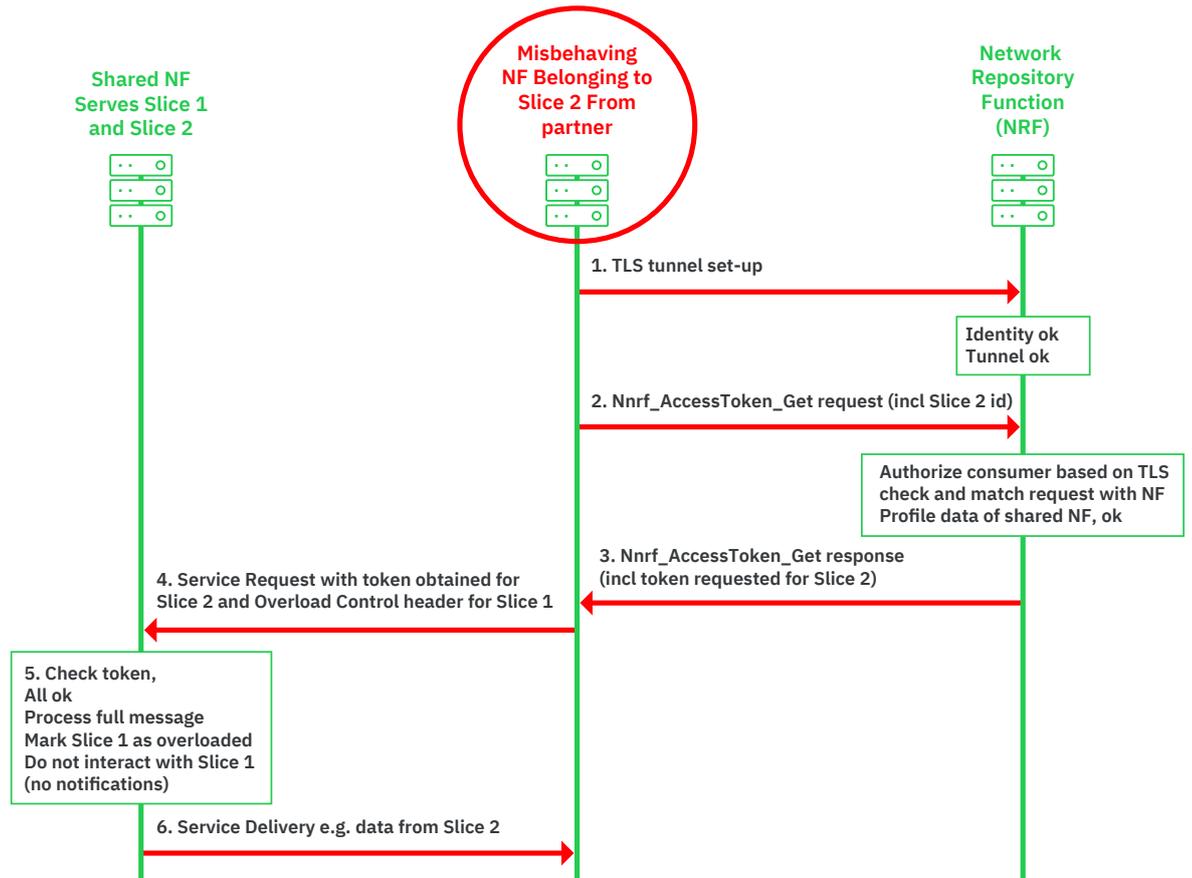


Figure 13: Potential DoS Attack Against Slice Network Functions

Currently, there is no requirement in the 3GPP specifications to validate if the slice identity in the 3GPP-Sbi-Oci header matches the slice identity in the token for the service API usage. Note, that the slice identity in the token is not clearly specified in 3GPP right now as it requires the usage of the additional scope field in the AuthenticationTokenClaims which is not defined in detail and hence would not provide interoperability between network functions from different vendors.

This kind of missing matching could potentially lead to misuse of the overload control features of 3GPP i.e. the OCI indicator, which could potentially result in partial network delays or outages.

There are of course many other possibilities to misconfigure a 5G network. We described many of those in our blogs on 5G Open Source Foundation for Application Security (OWASP) <sup>[9]</sup>, <sup>[10]</sup>, <sup>[11]</sup>. Configuration mistakes are not unusual in a complex system, therefore, often those systems are divided into groups of different sensitivity to protect key elements of the system in case of a breach. This brings us to the next challenge of missing security zones in 5G, which is an important topic as the key differentiator of 5G is the integration of third parties.

<sup>[9]</sup> Adaptive Mobile Security Blog, "5G Security Projections for the Future", 5. October 2020 <https://www.adaptivemobile.com/blog/5g-security-projections-for-the-future>

<sup>[10]</sup> Adaptive Mobile Security Blog, "Top 5G network security issues in 2028 - 5G OWASP for Networks", 19. October 2020, <https://www.adaptivemobile.com/blog/top-5g-network-security-issues-in-2028-5g-owasp-for-networks>

<sup>[11]</sup> Adaptive Mobile Security Blog, "OWASP: Top 5G Vertical Industry Security Risks in 2028", [https://www.adaptivemobile.com/blog/owasp-top-5g-security-risks-vertical-in-2028\\_2\\_november-2020](https://www.adaptivemobile.com/blog/owasp-top-5g-security-risks-vertical-in-2028_2_november-2020)



### 3.1.4 Security Challenges Through Missing Security Zones

We already raised the need for security zoning for the integration of legacy network nodes into the 5G core, but the topic of security zoning is much broader. Slicing brings new partners directly into the core network. For example, some slices belonging to an MVNO may have their own subscriber database functions and the MVNO has its own UDM and manages the UEs itself. The virtual instance of the UDM is then managed by the host MNO, but the subscriber data management is done by the MVNO.

This MVNO may also have its own UPF in its slice and provide local streaming e.g. at events. The MVNO would not generate the streaming content itself, but some application developers and content creators will do that for the MVNO. The MVNO slice network functions and host MNO network functions would have protection from roaming partners via the SEPP (see figure 14).

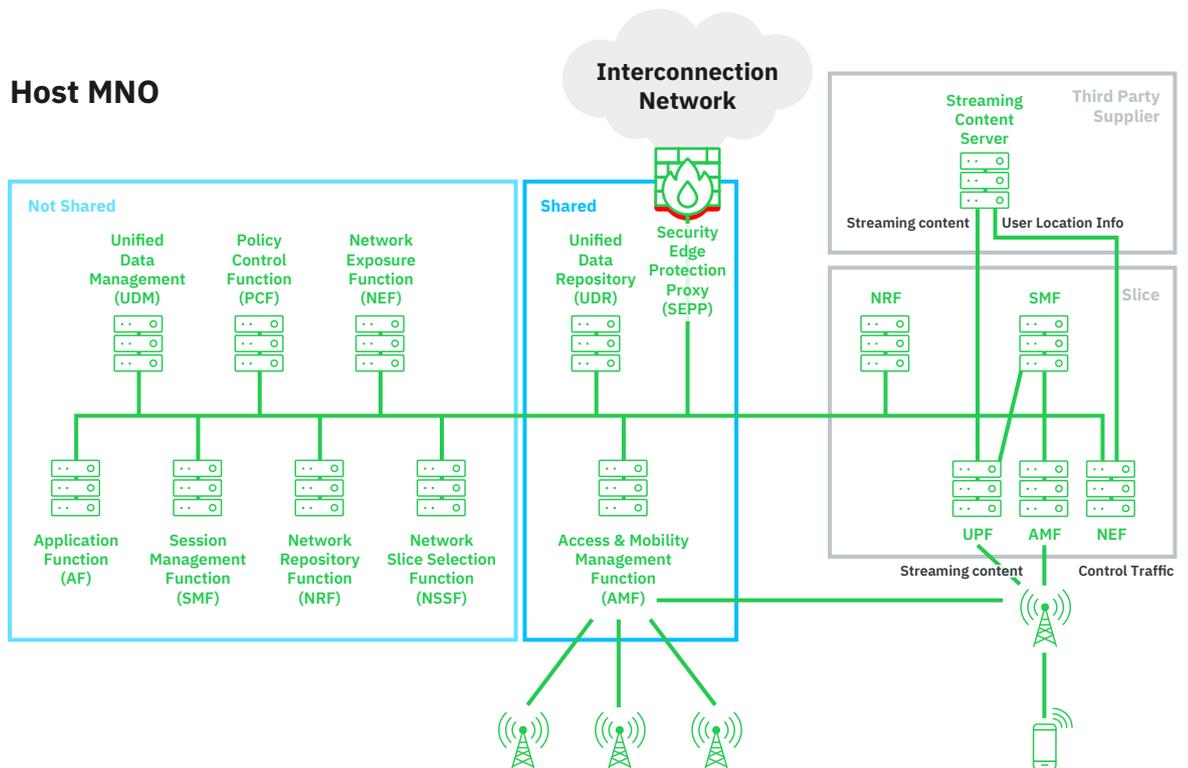


Figure 14: Third-Party Service for Slice Customer

For this kind of localized streaming service, the service is offered by the MVNO slice. The MVNO slice would need to know when the UE is in the proximity of the event to be able to provide the localized streaming service. Depending on slice configuration and setup, the location information of the UE could be:

- held by the AMF that is dedicated to the MVNO Slice, or
- be in the AMF that is shared with the host MNO, or
- not be present in AMF, and the location is requested from the UDM.

This means that the MVNO slice network functions would need to have access to the AMF or UDM in the MNO core to provide its service to the UEs.

This all depends strongly on the service type and the implementation, but in general it should be assumed that a slice customer of a hosting MNO has its ecosystem, suppliers, subcontractors etc. and the slice customer will bring these unique connections and business arrangement with them when they connect to the mobile core network of the hosting MNOs. This implies that a subcontractor or third-party gains access to core network functions via the MVNO slice. While the hosting MNO may have a good level of trust in the



MVNO, it may not have the same level of trust in a subcontractor of that MVNO. In the IT environment, we see attackers that move vertically through subcontractors to reach their targets<sup>[18]</sup>, so this risk may also be faced here. In general, by opening APIs and allowing external companies to control and access network functions in the core network, we must consider how to divide the network into different zones of trust.

Another security issue is inter-slice communication security. To illustrate the problem, we will describe the issue by using an example. Let's assume we have two slices in a network. A vehicular slice and an entertainment slice. A vehicular slice may interact with an entertainment slice for entertainment provisioning for the back-seat passengers. Therefore, one slice would communicate with another slice to provide the service. The UPF in the core network would provide the added entertainment content. The reason to have two separate slices could be that their QoS features are different. In our example, the vehicular slice would be a V2X and the other entertainment slice an ultra-reliable, low latency slice. The communication between those two slices would need to be authenticated, authorized on API level and filtered on information element level to avoid attacks like malicious traffic redirects or the addition of malware. For this, each slice would need to be contained in an individual security zone.

Another security zoning challenge relates to the SEPP. A normal standardized SEPP is not a network function in the 3GPP architecture as it does not have an SBI and offers services to the SBA. Therefore, if a message arrives at the SEPP coming from a slice in the home network, there is no interface for the SEPP to obtain information, if this slice is allowed to send messages outbound from the home network to the interconnection network. Also, the SEPP would not be able to validate if a network functions of a slice in the home network would be allowed to receive messages from visited networks. As the standard and not enhanced SEPP is not a 3GPP network function, it does not require an "authorization ticket" from the NRF to be used. Enhancing the SEPP to handle slice security is not yet covered by existing standards.

In this section we saw several examples where the classical "one core network/one trust zone" model that legacy networks have, does not work for 5G network and slicing. The SBA offers services via APIs to network functions, but the trust level of those network functions and the slices they host vary. This needs to be reflected in the architecture to avoid attacks.

### 3.1.5 Vulnerability: Security Challenges Through Missing Security Granularity

The SBA with all its flexibility and possibilities is challenging to configure correctly. As can be seen in the previous sections the slice identity (S-NSSAI) plays a crucial role in ensuring that the UEs receive the correct service and that the network is offering the service level that corresponds to that slice type. We also saw that the authorization in the SBA works on slice and service level. But what about the specific information elements e.g. cellID, UE identity etc. within the messages sent to and from slices?

In section 3.1.3 we discussed the issuing of the access token by the NRF and the related challenges related to trust in the presented slice identity. But there is another security challenge related to the granularity of authorization. The authorization token is bound to a service, network function and a slice identity, but the actual request to the service will contain all kinds of different information elements needed to provide the service.

Verification between information element and slice identity is important to stop a rogue slice from requesting information belonging to another slice. We already gave an example in 3.1.3, where the rogue slice could trick the NRF into issuing a token for another slice; now we will outline another attack vector on how to obtain information from another slice.

In this example in figure 15, the Location Information API description of the AMF allows the requesting of a UEs location. A rogue network function belonging to Slice 2 would set-up a TLS connection with the NRF in step 1. Then it sends a request to use the service of the shared-NF for Slice 2 (step 2). As the AMF is shared between Slice 1 and Slice 2, the NRF would generate in step 3 such a token and send it in step 4 to the rogue Network Function belonging to Slice 2. The messages in step 1 – step 4 are legitimate and valid normal

<sup>[18]</sup> Jessica Davis, Health IT Security "50% of Cyberattacks Target Supply Chain, Seek Lateral Movement", 3. April 2019, <https://healthitsecurity.com/news/50-of-cyberattacks-target-supply-chain-seek-lateral-movement>



operational messages. The rogue network function then creates a request to the shared NF that provides the service that contains the token for its own Slice 2 (we assume again that the optional additional scope element in the AuthenticationTokenClaim is used for the slice identity), but the UE identity in the actual service request is from the other Slice 1 (step 5). The shared-NF would assume that the NRF has performed all authorization duties, it checks that the token is valid. However, this validity check simply verifies the correctness of the NRF issued token for Slice 2 and if the request is coming from the correct Slice 2 belonging network function. The shared NF does not validate in step 6 if the UE identity also belongs to Slice 2, therefore, it would assume that the request is valid and deliver the location of the user in step 7.

During the whole process, there was no check if the UE identity in the service request sent from the rogue network function to the shared network function should be accessible to Slice 2.

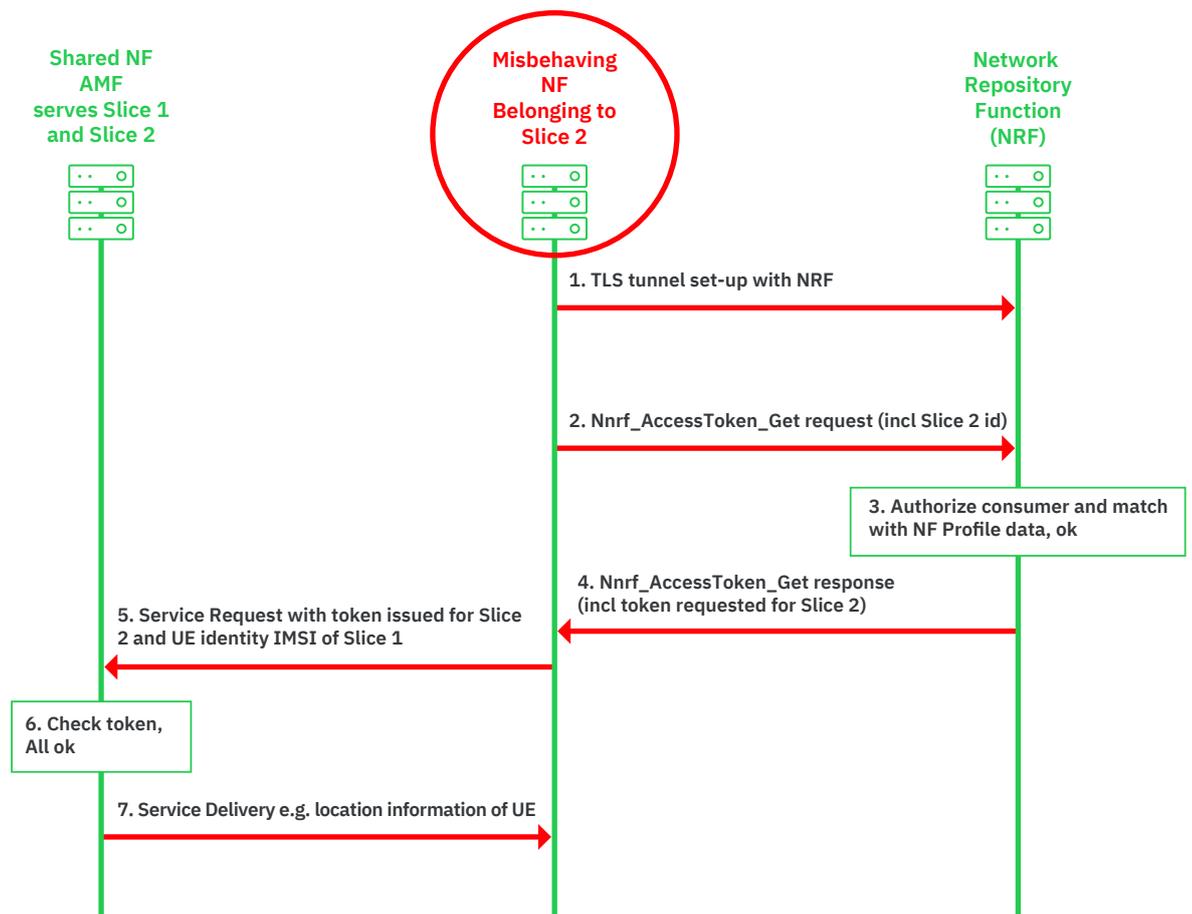


Figure 15: Risk of Information leakage Between Shared and Dedicated NFs via Information Elements

In this section we have seen several security challenges for 5G slicing related to missing inter-layer checks, configuration granularity or missing security zones. We will now recommend countermeasures to tackle those issues and provide robust network security that can support the integration of third parties and services.



### 3.2 Security Countermeasures and New Features

In the previous sections we saw many challenges related to interworking, complexity, inter-layer matching, configuration, security zones and filtering granularity. Some of them are solvable to some degree by amending the corresponding specifications and patching existing network functions after the specifications are updated, however, many will need the introduction of new security measures. The expected amount of complexity, new partners and opening-up of the network makes it necessary to split the network into clear security zones on the signalling layer.

We saw potential attack paths that exploit the lack of cross-validation between the layers. Therefore, it is important to bind the layers together in a security holistic approach that encompasses the signalling layer. A pure IP layer firewall or general transport layer security solution cannot provide such a holistic approach, as it does not have the understanding of the interaction of the layers, such as whether the slice identity in the actual signalling layer request matches the transport layer, or if a UE identity belongs to a slice or not. Therefore, the deployment of an IP firewall gives a false sense of security, as the controls provided by it can be bypassed on the signalling layer.

From our research, based on the analysis of the challenges we see, the following key point for applying signalling security filters could be installed between zones:

- Between roaming partners and hosting network
- Between not-shared and shared network functions
- Between third party and shared network functions
- Between 5G and legacy network elements
- Between the core network and external partners
- For inter-slice communication between slices

We illustrate the different approaches in separate (figures 16 and 17) for readability. The first step is to group network functions that are dedicated exclusively to a particular slice. If inter-slice communication is required, then that link between those slices must be secured and filtered on the signalling layer. Then there are network functions (NFs), which are shared between the slices and potentially the host MNO. Those NFs are potentially vulnerable as outlined in our challenges section and would need to be protected on the signalling and information element level from those attacks through fine-grained filtering. This zone would also belong to the SEPP as some slices may require roaming support. The SEPP would offer protection of the slices from attacks coming from the roaming network.

The network functions which are dedicated to the host MNO and shared with slices would require their own security zone which would need to be protected on the signalling layer from malicious requests.



Below is figure 16 for the case of different slices:

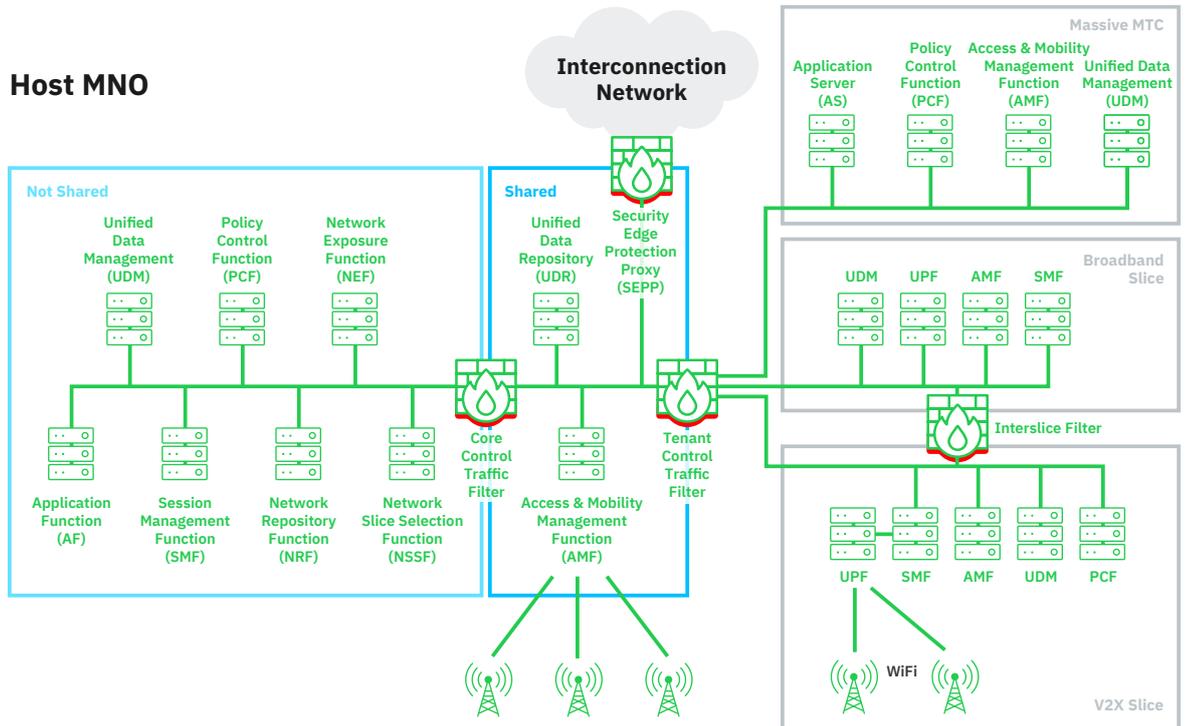


Figure 16: Example Security Zoning in Sliced Network

These kinds of security zones provide security for slices in the figure above, but it is important to remember that these slices might also have interfaces to external parties like content suppliers. Basically, in addition to the protections above, we would need to secure the communication on all external 3rd parties. For those kinds of 3rd party interfaces, we recommend the following security and protection points:

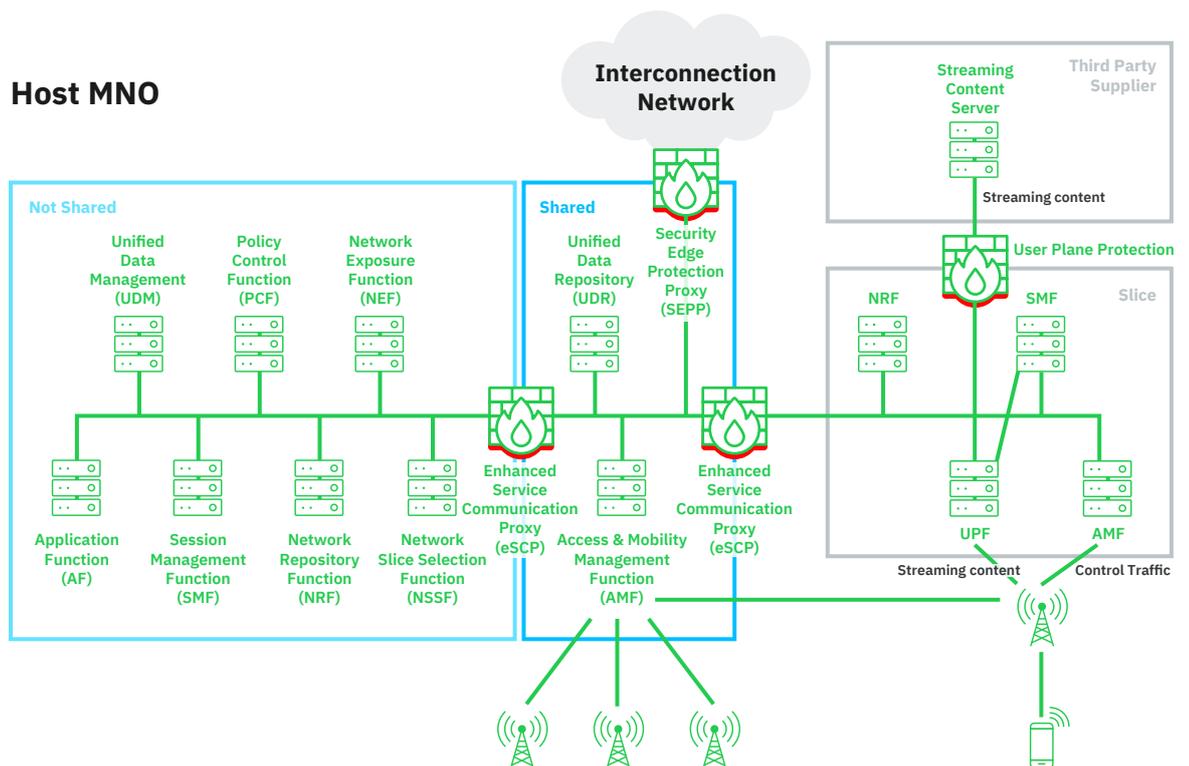


Figure 17: Full Network Protection



In figure 17 a third party supplies the slice with content. This incoming user plane traffic from the content server would need to be filtered for malicious content by a user plane protection filter. We discussed migration protection for signalling traffic in full detail in our previous white paper <sup>[6]</sup>.

On an architectural level, we now understand how the network can be divided into security zones and where signalling security protection can be placed to create those security zones. The next step is to understand the features a signalling security protection solution needs to have to counter the potential attack vectors outlined in section 3.1.3 and 3.1.5,

In the challenges, we saw that being authorized to use a service on a network function using the NRF is not sufficient to protect against data leakage. The resource holding node might be tricked into revealing information from another slice (section 3.1.5) or the NRF may issue an access token for a different slice (section 3.1.3). Protecting third party access to the core network (section 3.1.4) also poses a challenge in terms of filtering and access granularity.

Another attack angle is to exploit the missing correlation between layers. Information cross-correlation between layers prevents impersonation and spoofing attacks. We successfully use such layer correlation for SS7 and Diameter attacks. This kind of correlation can be done at the protection nodes described in figure 16 and figure 17.

Many suspicious signalling “attacks” that we see today in the core network are not malicious at all, but just the results of incorrect configurations. In the future, when we see a much more complex network and many 3rd party slices, it is expected that one or more of these slices globally will either send non-standard messages, be misconfigured leading to unexpected behaviour, or exhibit misbehaviour for no reason other than being the slice used by strangely behaving IoT devices. A load in one slice should not impact the rest of the network or other slices, even if nodes are shared. These kinds of load-related attacks need to be detected and prevented by the protection nodes. Therefore, any kind of signalling layer protection requires deep knowledge of the telecommunication API information elements and protocol mechanisms to prevent an attacker bypassing filtering and control mechanisms. Such signalling protection solutions can be deployed as dedicated protection nodes or protection network functions.

We now have a range of requirements for potential protection nodes to protect within and between slices, and 3rd party APIs.

3GPP does not recommend such a node right now, but there is the Service Communication Proxy, which could be potentially enhanced to support those requirements above and provide the baseline for such protection solutions. Those extensions would go way beyond its current scope. These extended SCPs would reside at the edges of the security zones and protect against the attacks outlined in section 2. There it would protect the network by:

- Validating the correctness of message formats
- Enforcing service level agreement on information element, slice and node level
- Correlating information between layers and protocols
- Providing load-related functionality to prevent DoS attacks

Also, when acting as a proxy towards legacy nodes and protocols it can provide interworking functionality and protection against attacks coming from legacy nodes.

Within AdaptiveMobile, we define this as an enhanced SCP (eSCP), as it goes beyond what 3GPP has specified and provides additional filtering and protection features. Each security zone protection node will need its own type of protection depending if it interacts with third parties, legacy network nodes, slices or roaming partners.

The different protection nodes in the network can correlate their attack information to enhance efficiency for an attack using multiple entry vectors. We observed these kinds of multi-entry vector attacks in our threat analysis <sup>[12]</sup> and expect that attackers will also utilize different approaches in 5G to obtain the desired information.



## 4. Summary and Conclusion

While the access from the mobile phone user to a specific slice is well-secured, the core network aspects of who has access to slice specific information are not, and often left to individual deployment decisions and configurations. We outlined several ways, within the slicing model, how information might potentially be exposed, how services could be misused and how Denial of Service attacks could be executed against network elements. The usage of IP firewalls and usage of TLS does not provide complete protection against impersonation. IP firewalls only provide integrity and confidentiality for communication on lower layers between network functions, but do not protect attacks from partners or network functions that go “rogue”. Also, fine-grained security controls are needed to ensure that no data is leaked, and service level agreements can be enforced on the information element level. These attacks were studied in the context of a single core network and a misbehaving network function, but the same attack vector approach can also potentially be applied using messages over the interconnection network.

Authentication and authorization on slice level exist, but there are still many open security issues for

- Protection against misconfigured, misbehaving or compromised slices
- Information extraction by a rogue slice from another slice e.g. via shared elements
- Interslice communication security
- Roaming and slicing security on the information element level
- Interaction with legacy network nodes
- API Security with external 3rd parties
- Cross-validation between layers
- Cross-correlation of attacks between different protocols

In 5G, mobile networks will be open to new partners, but those partners bring new risks to the core network. Our recommendation to deal with these risks is to use an enhanced filtering and validation approach, which combines information from different layers, protocols and integrates external threat information. This kind of filtering and validation approach allows division of the network into security zones and safeguarding of the 5G core network. Cross-correlation of attack information between those security network functions maximizes the protection against sophisticated attackers and allows better mitigations and faster detection while minimizing false alarms. This enhanced security network functions should provide the operational staff in the MNO with a unified holistic view of the security state of the network and easy integration of the latest threat intelligence, meaning that no matter how complex the network gets through slicing, subscribers and the underlying networks are safe.



## Abbreviations

<b>2G</b>	Second Generation
<b>3G</b>	Third Generation
<b>3GPP</b>	3rd Generation Partnership Project
<b>4G</b>	Fourth Generation
<b>5G</b>	Fifth Generation
<b>5GS</b>	5G System
<b>AAA</b>	Authentication, Authorization and Accounting
<b>AAA-S</b>	Authentication, Authorization and Accounting Server
<b>AMF</b>	Access and Mobility Management Function
<b>API</b>	Application Programming Interface
<b>APN</b>	Access Point Name
<b>AS</b>	Application Server
<b>CellId</b>	Cell Identity
<b>DNN</b>	Data Network Name
<b>DoS</b>	Denial of Service
<b>EAP</b>	Extensible Authentication Protocol
<b>eMBB</b>	enhanced Mobile Broadband
<b>eSCP</b>	enhanced Service Communication Proxy
<b>GPRS</b>	General Packet Radio Service
<b>GSM</b>	Global System for Mobile Communications
<b>GSMA</b>	GSM Association
<b>GTP</b>	GPRS Tunneling Protocol
<b>GTP-C</b>	GTP – Control Plane
<b>GTP-U</b>	GTP – User Plane
<b>GST</b>	Generic Slice Template
<b>hNRF</b>	home Network Repository Function
<b>HSS</b>	home Subscriber Server
<b>HTTP</b>	Hypertext Transfer Protocol
<b>hSEPP</b>	Home Security Edge Protection Proxy
<b>IE</b>	Information Element
<b>IoT</b>	Internet of Things
<b>IP</b>	Internet Protocol
<b>IPSec</b>	IP Security
<b>IPX</b>	Interconnection Network
<b>IT</b>	Information Technology
<b>mMTC</b>	massive Machine Type Communication
<b>MNO</b>	Mobile Network Operator
<b>MVNO</b>	Mobile Virtual Network Operator
<b>NAS</b>	Non-Access Stratum
<b>NEF</b>	Network Exposure Function
<b>NEST</b>	NEtwork Slice Type
<b>NF</b>	Network Function
<b>NFProfile</b>	Network Function Profile

<b>NG</b>	Network Group
<b>NR</b>	New Radio
<b>NRF</b>	Network Repository Function
<b>NSSAA</b>	Network Slice Specific Authentication & Authorization
<b>NSSAAF</b>	Network Slice Specific Authentication & Authorization Function
<b>NSSAI</b>	Network Slice Selection Assistance Information
<b>NSSF</b>	Network Slice Selection Function
<b>OAuth</b>	Open Authorization
<b>OCI</b>	Overload Control Indicator
<b>OSI</b>	Open Systems Interconnection
<b>OWASP</b>	Open Source Foundation for Application Security
<b>PCG</b>	Policy Control Function
<b>PDU</b>	Packet Data Unit
<b>PLMN</b>	Public Land Mobile Network
<b>QoS</b>	Quality of Service
<b>RAN</b>	Radio Access Network
<b>RAT</b>	Radio Access Technology
<b>REST</b>	Representational State Transfer
<b>SBA</b>	Service Based Architecture
<b>SBI</b>	Service Based Interface
<b>SCP</b>	Service Communication Proxy
<b>SD</b>	Slice Differentiator
<b>SEPP</b>	Security Edge Protection Proxy
<b>SLA</b>	Service Level Agreement
<b>SMF</b>	Session Management Function
<b>SMS</b>	Short Message Service
<b>S-NSSAI</b>	Single-Network Slice Selection Assistance Information
<b>SS7</b>	Signalling System No 7
<b>SST</b>	Slice Service Type
<b>TLS</b>	Transport Layer Security
<b>TR</b>	Technical Report
<b>TS</b>	Technical Specification
<b>UDM</b>	Unified Data Management
<b>UDR</b>	Unified Data Repository
<b>UE</b>	User Equipment
<b>URLLC</b>	Ultra-Reliable Low Latency Communications
<b>UPF</b>	User Plane Function
<b>V2X</b>	Vehicle to X
<b>vNRF</b>	Visited Network Repository Function
<b>vSEPP</b>	Visited Security Edge Protection Proxy
<b>WiFi</b>	Wireless Fidelity



## References

- <sup>[1]</sup> GSMA FS.36, “5G Interconnect Security”, version 1.0, 20. November 2020
- <sup>[2]</sup> 3GPP TS 23.501, “System architecture for the 5G System (5GS)”, version 16.6.0, 24. September 2020
- <sup>[3]</sup> GSMA NG.116, “Generic Network Slice Template”, version 3.0, 22. May 2020
- <sup>[4]</sup> 3GPP TR 28.801, “Telecommunication management; Study on management and orchestration of network slicing for next generation network”, version 15.1.0, 4. January 2018
- <sup>[5]</sup> 3GPP TS 28.531, “Management and orchestration; Provisioning”, version 16.7.0, 25. September 2020
- <sup>[6]</sup> Adaptive Mobile Security, Silke Holtmanns, Cathal McDaid, “Securing the Path from 4G to 5G: How to protect mobile networks and subscribers during migration”, <https://www.adaptivemobile.com/downloads/securing-the-path-from-4g-to-5g>
- <sup>[7]</sup> 3GPP TS 29.510, “5G System; Network function repository services; Stage 3”, version 16.5.0, 25. September 2020
- <sup>[8]</sup> 3GPP TS 29.500, “5G System; Technical Realization of Service Based Architecture; Stage 3”, version 16.5.0, 24. September 2020
- <sup>[9]</sup> Adaptive Mobile Security Blog, “5G Security Projections for the Future”, 5. October 2020 <https://www.adaptivemobile.com/blog/5g-security-projections-for-the-future>
- <sup>[10]</sup> Adaptive Mobile Security Blog, “Top 5G network security issues in 2028 - 5G OWASP for Networks”, 19. October 2020, <https://www.adaptivemobile.com/blog/top-5g-network-security-issues-in-2028-5g-owasp-for-networks>
- <sup>[11]</sup> Adaptive Mobile Security Blog, “OWASP: Top 5G Vertical Industry Security Risks in 2028”, <https://www.adaptivemobile.com/blog/owasp-top-5g-security-risks-vertical-in-2028>, 2. November 2020
- <sup>[12]</sup> Adaptive Mobile Security, “SIMjacker White Paper”, <https://simjacker.com/>
- <sup>[13]</sup> 3GPP TS 23.502, “Procedures for the 5G System (5GS)”, version 16.6.0, 24. September 2020
- <sup>[14]</sup> 3GPP TS 33.501, “Security architecture and procedures for 5G System”, version 16.5.0, 16. December 2020
- <sup>[15]</sup> GSMA NG.113, “5G Roaming Guidelines”, version 2.0, 28. May 2020
- <sup>[16]</sup> 3GPP TS 29.531, “5G System; Network Slice Selection Services; Stage 3”, version 16.4.0, 25. September 2020
- <sup>[17]</sup> Fierce Wireless, “‘Real’ 5G relies on 5G NR, Standalone architecture: Special Report”, 24. May 2020, <https://www.fiercewireless.com/wireless/real-5g-relies-5g-nr-standalone-architecture-special-report>
- <sup>[18]</sup> Jessica Davis, Health IT Security “50% of Cyberattacks Target Supply Chain, Seek Lateral Movement”, 3. April 2019, <https://healthitsecurity.com/news/50-of-cyberattacks-target-supply-chain-seek-lateral-movement>



## Annex: Timeline/Responsible Disclosure Process

AdaptiveMobile Security discovered these attacks at the end of 2020 and submitted the three attacks as CVDs to the GSMA on the 4th of February 2021. These attacks were recognized by GSMA as vulnerabilities and given the designation CVD-2021-0047. As part of the mitigation strategy GSMA have identified the affected GSMA documents and is currently reviewing potential improvements to their own recommendations. In addition, GSMA is taking steps to interact with 3GPP to bring the attacks and potential countermeasures to their attention. The full mitigation of the described attacks may require some larger feature development in 3GPP specifications and could potentially be part of Release 17. As part of its responsible disclosure process AdaptiveMobile Security has been assisting the GSMA and other interested parties throughout the disclosure period and will continue to do so in the future.

## About AdaptiveMobile Security

AdaptiveMobile Security is a world leader in mobile network security, everyday protecting over 80 Mobile Operators and billions of mobile subscribers and devices globally from fraudsters, criminals and nation states. We have the strongest 5G core network security team, who are designing, planning and building the very best in 5G core network security solutions focussing on threat-intelligence, security heritage and protocol correlation.

AdaptiveMobile Security brings a unique security perspective on real-time mobile network traffic. The global insight provided by our 5G, Signalling and Messaging thought leaders, security specialist teams and Threat Intelligence Unit, combined with our signalling and network protection software that sits at the heart of the network, ensures AdaptiveMobile Security remains at the forefront of the latest advancements in mobile networks and their security, and continues to be the trusted partner of many of the world's largest Mobile Operators.

For more information on how AdaptiveMobile Security can help you migrate to a secure 5G core network please contact [sales@adaptivemobile.com](mailto:sales@adaptivemobile.com)

## Legal Notices

© 2021 AdaptiveMobile. All rights reserved. This document, or any part thereof, may not, without the written consent of Adaptive Mobile Security Limited, be copied, reprinted or reproduced in any material form including but not limited to photocopying, transcribing, transmitting or storing it in any medium or translating it into any language, in any form or by any means, be it electronic, mechanical, optical, magnetic or otherwise.

AdaptiveMobile, Network Protection Platform, and Policy Filter are trademarks of Adaptive Mobile Security Ltd.

All other products are trademarks or registered trademarks of their respective owners and are hereby recognised as such.

The information contained herein is believed to be accurate and reliable. Adaptive Mobile Security Ltd. accepts no responsibility for its use by any means or in any way whatsoever. Adaptive Mobile Security Ltd. shall not be liable for any expenses, costs or damage that may result from the use of the information contained within this document. The information contained herein is subject to change without notice.

### Head Office

Ferry House, 48-52 Lower Mount St, Dublin 2.

Contact: [sales@adaptivemobile.com](mailto:sales@adaptivemobile.com)

[www.adaptivemobile.com](http://www.adaptivemobile.com)

### Regional Sales Contact Numbers

US, Canada, Latin America Sales: +1 972 377 0014

UK Sales: +44 207 049 0421

Middle East Sales: +97144 33 75 83

Africa Sales: +27 87 5502315

Asia Sales: +65 31 58 12 83

European Sales: +353 1 524 9000

### Regional Operational Support Contact Numbers

UK: +44 208 584 0041

Ireland: +353 1 514 3945

India: 000-800-100-7129

US, Canada: +1 877 267 0444

LATAM: +525584211344

