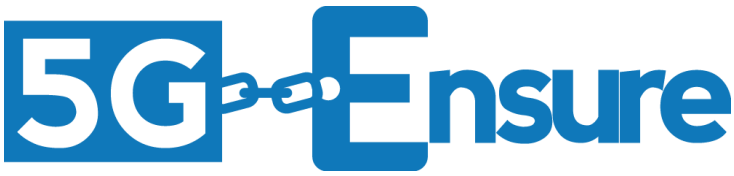

Deliverable D2.4
Security Architecture (draft)

Project name	5G Enablers for Network and System Security and Resilience	
Short name	5G-ENSURE	
Grant agreement	671562	
Call	H2020-ICT-2014-2	
Delivery date	31.10.2016	
Dissemination Level:	Public	
Lead beneficiary	Ericsson AB (EAB)	Mats Näslund, mats.naslund@ericsson.com
Authors	EAB: Alexander Maximov, Mats Näslund, Per Ståhl IT Innov: Gianluca Correndo, Vadims Krivcovs, Stephen Philips LMF: Vesa Lehtovirta, Vesa Torvinen NEC: Felix Klaedtke Nixu: Seppo Heikkinen, Tommi Pernila Nokia: Hon-Yon Lach, Linas Maknavicius Orange: Ghada Arfaoui, Jean-Philippe Wary Oxford: Ravishankar Borgaonkar, Piers O'Hanlon SICS: Rolf Blom, Martin Svensson TCS: Sébastien Keller TS: Edith Felix, Pascal Bisson VTT: Petteri Manersalo, Pekka Ruuska, Jani Suomalainen, Janne Vehkaperä	

Executive summary

This deliverable (D2.4) of the 5G-ENSURE project describes a draft security architecture for 5G networks. The focus lies on a logical and functional architecture and omits (most) aspects related to physical/deployment architecture. This focus is motivated by general trends such as network de-perimetrization as well as 5G systems' strong dependency on software defined networking and virtualization in general. Furthermore, this focus has reduced the otherwise strong interdependency between this architecture task and the trust modelling and risk analysis tasks in 5G-ENSURE. Still, each of these three tasks have at the time of writing produced initial draft documents, which will then be re-used in a second iteration of all three tasks, producing updated, final versions.

The project's 5G security architecture builds on, extends (and in our opinion clarifies) the current 3GPP security architecture. The *logical* "dimension" of our architecture captures first of all security aspects associated with the various *domains* that are involved in delivering services over 5G networks. This part is therefore also strongly associated with the project's trust model. Additionally, the logical part captures security aspects associated with network layers and/or special types of network traffic. This is in our architecture associated with different *strata*. The *functional* "dimension" of our architecture comprises a set of security capabilities required to protect and uphold the security of the various domains and strata. In the functional dimension, we build on the 3GPP defined *security feature groups*. We also here extend and refine to adapt to a 5G context.

A goal of the architecture work within 5G-ENSURE has been to clearly provide rationale for the architecture's structure and features, i.e. instead of staring from detailed security requirements, we seek to motivate which high level *security problem* is relevant in a 5G context, and then break that down into a manageable set of *security objectives* for 5G. From these objectives, the high level architecture is derived, and only after that stage detailed requirements enter the discussion (many of them defined in Task 2.3: Risk assessment, mitigation and requirements). Conversely, care has been taken to provide means for performance indicators such as *measurability*, i.e. simplifying the task of validating that the proposed architecture and its features meet the objectives, and, that the objectives appropriately address the security problem. This last aspect is not covered to any depth in this first draft.

Foreword

5G-ENSURE belongs to the first group of EU-funded projects which collaboratively develop 5G under the umbrella of the 5G Infrastructure Public Private Partnership (5G-PPP) in the Horizon 2020 Programme. The overall goal of 5G-ENSURE is to deliver strategic impact across technology and business enablement, standardisation, and vision for a secure, resilient, and viable 5G network. The project covers research & innovation - from technical solutions (5G security architecture and testbed with 5G security enablers) to market validation and stakeholders' engagement - spanning various application domains.

This deliverable, "Security Architecture (draft)" (D2.4), is a first approximation of an overarching security architecture, suitable for 5G networks, produced by task T2.4 in work package WP2 "Security requirements and architecture" of 5G-ENSURE. The reason this document can only be a first approximation is due mainly to the *strong* interdependency on the tasks T2.1, T2.2, and T2.3 but also due to the need for coordination within the 5G-PPP. While T2.1 (use cases) was completed before this work started, T2.2 (Trust model) and T2.3 (Risk analysis and requirements) have been running in parallel. In addition, T2.2 and T2.3 are dependent on the results of T2.4. Therefore, it has been necessary to structure the work with draft deliverables from tasks T2.2, T2.3, and T2.4 amid the running project. A second iteration in all three tasks, with mutual leverage from each other's draft deliverables, will result in updated, final reports at the end of the 5G-ENSURE project.

Despite the draft status of this deliverable, we believe that the focus on logical and functional aspects will provide useful guidance for creating a shared 5G security architecture vision within the 5G-PPP as well as providing a useful basis for ongoing standardization discussions.

Disclaimer

Copyright notice

Contents

Deliverable D2.4 1
Security Architecture (draft) 1
Foreword 3
Disclaimer 3
Copyright notice 3
Contents 4
1 Introduction 8
1.1 Why a Security Architecture? 9
1.2 Why a New Security Architecture? 10
1.3 5G-ENSURE Methodology 10
1.4 Abbreviations 11
2 The Security Landscape of 5G 12
2.1 Security Characteristics of 5G Capabilities 12
2.1.1 Virtualisation 12
2.1.2 Network Slicing 13
2.1.3 Mobile Edge Computing 14
2.1.4 Internet of Things 15
2.1.5 Heterogeneous Network Access 15
2.1.6 Critical Service Support 16
2.2 Business and Trust Model Summary 16
2.3 High Level Threats to 5G Systems 17
2.3.1 Smart Jamming Attacks 17
2.3.2 Radio Access Network Threats 18
2.3.3 SDN/NFV Related Threats 18
2.3.4 Location Based Attacks 18
2.3.5 Small Cells Security 19
2.3.6 Cyber Security Threats 19
2.4 Regulatory Aspects 19
2.5 The 5G Security Problem Definition 21
3 5G Security Objectives 22
3.1 General 22
3.2 Security-relation to Legacy Systems 22
3.3 RAN/RAT Security 22

3.4 Secure Virtualization 22
3.4.1 General 22
3.4.2 Slicing 23
3.5 IoT Security 23
3.6 Security Management 23
3.7 New Business and Use-cases 23
3.8 Regulatory Aspects 24
4 5G Security Architecture Overview 24
4.1 Rationale 24
4.2 Domains 25
4.2.1 Domain Types 28
4.2.2 Infrastructure Domains 28
4.2.3 Tenant Domains 28
4.2.4 (Additional) UE Domain 29
4.2.5 Compound Domains 30
4.3 Strata 31
4.3.1 Management Stratum 33
4.4 Security Feature Groups 33
5 Architecture Enforcement 35
5.1 Introduction 35
5.2 5G Domain Security Architecture Enforcement 36
5.2.1 Infrastructure Domains 36
5.2.2 Tenant Domain 37
5.2.3 (Additional) UE Domain 39
5.2.4 Compound Domains 39
5.2.5 Domain Interactions 39
5.3 5G Strata Security Architecture Enforcement 41
5.3.1 Application Stratum 41
5.3.2 Home Stratum 41
5.3.3 Serving Stratum 41
5.3.4 Transport Stratum 41
5.3.5 Management Stratum 41
6 Trust Model Mapping 41
6.1 Relationship Between Business Models, Trust and Security 42

6.1.1 Relationship to Architecture 43
6.1.2 Stakeholder and Trust Models 44
6.2 Mapping of Actors to Architectural Domains 46
6.2.1 Analysis 47
7 5G Security Design Principles and Recommendations 47
7.1 Security Concepts 48
7.1.1 Authentication 48
7.1.2 Confidentiality 49
7.1.3 Integrity 49
7.1.4 Availability 49
7.1.5 Non-repudiation 50
7.1.6 Secure Interworking 50
7.2 Standards Based Security 50
7.2.1 ETSI 50
7.2.2 NIST 50
7.2.3 IETF 50
7.2.4 ITU-T 51
7.2.5 ISO 51
7.3 Further Recommendations 51
7.3.1 Implementing Security 51
7.3.2 Design Phases 52
7.3.3 Monitoring 53
7.3.4 Orchestration 54
8 Mapping of 5G-ENSURE Security Enablers 55
8.1 Security Monitoring Enablers 56
8.1.1 Analysis of the Security Monitoring Enablers 56
8.1.2 Mapping to Domains 58
8.1.3 Mapping to Strata 59
8.1.4 Mapping to Security Features Group 59
8.2 Privacy Enablers 60
8.2.1 Mapping to Domains 60
8.2.2 Mapping to Strata 60
8.2.3 Mapping to Security Features Group 61
9 Existing Work 61

9.1 3GPP 62
9.2 ITU X.805 63
10 Quality Attributes of the Architecture 64
11 Summary and Conclusions 64
12 References 66

1 Introduction

This deliverable, produced by the partners of the 5G-ENSURE project, describes a draft security architecture for 5G networks. The focus lies on a logical and functional architecture and omits (most) aspects related to physical/deployment architectures. This is motivated by aspects such as network de-perimetrization and 5G systems' strong dependency on software defined networking and virtualization in general. These aspects make the question "Who/what to protect?" more important to firstly consider in the more logically and functionally abstract sense than in the physical sense. The question "Where to place the protection, physically?" can in some situations only be answered by "Wherever the who/what is located". That is, security gets even more strongly coupled to the abstract assets themselves. Thus, relying on protection from the physical environment ("inside site A") or physical network topology ("behind firewall F"), is no longer meaningful. This does however not mean that physical placement of security features becomes irrelevant. Indeed, some dominating features of the current 3G and 4G security architectures have been driven by, for example, threats to radio base stations in physically exposed locations, tampering threats to user credentials in devices, and so on. Although these aspects will certainly continue to play a role in 5G, they are in a sense by now well understood and can therefore be given secondary priority. There will however also be new physical aspects of security entering in 5G. For example, the increased need for "hardware root of trust" in virtualized settings.

Another reason for focusing in the security architecture work on the logical/functional dimensions is due to a strong interdependency between some tasks in 5G-ENSURE, in particular task T2.4 (producing this report) and the tasks T2.2 and T2.3 (delivering trust model and risk analysis/requirements). Each of these three tasks have at the time of writing produced draft documents, which will then be re-used in a second iteration of all three tasks, producing updated, final versions. Architecture level coordination within the 5G-PPP is also a necessity that is under way. Prioritizing a logical and functional architecture have made the inter-task dependencies better manageable.

In the original description of work of the 5G-ENSURE project, the proposed working hypothesis was to build on the ITU-T X.805 security architecture, [x805]. However, after a finer analysis performed in the initial work of this task, some issues surfaced that turned out to make X.805 less attractive for our purposes. These issues are discussed in Section 9.2. Section 9 will also elaborate why it was instead chosen to build (mainly) on the 3GPP architectures, [ts33.102, ts33.401]. Consequently, the security architecture herein presented can be seen as extending the current 3GPP security architecture to a 5G context. In our opinion, the new architecture also resolves some lack of clarity in the 3GPP architecture.

Thus, borrowing from 3GPP, the *logical* part of our security architecture captures first of all security aspects associated with the various *domains* that are involved in delivering services over 5G networks. As many of the domains will typically be coupled to administration/ownership, these domains are therefore also strongly connected to the 5G trust model (Task T2.2). Additionally, the logical part captures security aspects associated with special types of network traffic and/or protocol layers. This is in our architecture associated with different *strata*. Finally, the *functional* dimension comprises the set of security capabilities required to protect and uphold the security of the various domains and strata. We here build on the 3GPP defined *security feature groups*, but again extend and also add a more fine grained division.

An additional goal of the architecture work has been to clearly provide rationale for the architecture's structure and specific features, i.e. instead of starting from detailed security requirements, we start out from the view point of a high level *security problem* that characterises the overall "5G security landscape" in a

"nutshell". We then break down this security problem into a manageable set of *security objectives* for 5G. Only after that stage detailed requirements enter the discussion (mainly those defined by the task T2.3). Conversely, care has been taken to provide means for *measurability*, i.e. simplifying the task of validating that the architecture and its features meet the objectives, and, that the objectives appropriately address the security problem. Readers who are familiar with Common Criteria [cc] (CC) may recognize aspects of this structure from the standardized format of CC assurance documentation. It should be noted though that we make no claims to have produced a complete "CC Protection Profile" for the overall 5G context.

Before moving over to the aforementioned security problem definition, the security objectives and, ultimately, the presentation of the architecture, we first provide some general rationale for producing "yet another security architecture".

1.1 Why a Security Architecture?

Many technical studies, in particular in the ICT and telecom area feel obliged to include an overall architecture description. The need for an architecture can be driven by the desire to produce high-level overviews, show completeness and/or soundness of a design, etc. A security architecture is on this level no different, i.e., it most often serves the same type of purposes. Some important differences and issues however become visible on closer consideration.

First of all, the security architecture is typically produced *for* something else, i.e. security is always there to support something, it is not a self-purpose. This observation is not really new, but unfortunately it is still not always well handled. One can find examples of architectures, showing figures with great detail in e.g. the network functions and their interfaces. Then, looking in the upper left corner of the figure, one finds a "box" labelled "security". Not only is the handling inadequate, one can simply never *understand* security by "boxing it in". It is tempting to draw parallels to research in philosophy/neurology trying to understand human consciousness. The French philosopher and mathematician Descartes believed that human consciousness was located (or at least centred) at one specific place in the brain (the pineal gland). Though science is still far from understanding human consciousness, it has at least been recognized that consciousness is more of an architectural aspect of the brain no attempt to model it as a "box" can lead to an understanding.

Indeed, the interplay between a system/network architecture and an associated security architecture can have far reaching consequences. A poorly designed network architecture can make it impossible (or at least very expensive) to adequately protect it. Conversely, a poorly designed security architecture may be highly secure, but at the same time negatively impact the network services (in extreme cases, even prevent them). Thus, the need to find the correct trade-offs between security and other requirements (availability, performance and functionality) is essential for the success of any system. Software defined networking can offer better possibilities here, offering flexibility without sacrifice. The point we wish to make is that the cross-coordination on security within the 5G-PPP will be crucial for the overall success of the initiative, and that it needs to start early. A 5G-PPP Security Working Group has been established and one of its first priorities is to coordinate the security work. 5G-ENSURE is also particioating in the 5G-PPP Architecture WG.

In the *mobile network setting*, the security architecture also plays a distinguished role in that, due to the regulatory aspects of telecommunications and spectrum usage, many aspects of the security architecture are strongly correlated with ability to fulfil contractual and legal obligations, to settle liabilities, etc.

1.2 Why a New Security Architecture?

The current 3G/4G networks already have security architectures defined in [ts33.102,ts33.401]. An analysis of the existing 3GPP architectures will be given in Sections 4 and 9. Let us begin by just summarizing the most important reasons why we simply cannot re-use them as-is and why a new security architecture for 5G is needed.

Trust model: There is no explicit and complete trust model documented for 3G and 4G networks. This does not imply in any way that they are insecure for their designed purpose, but it does produce an issue in 5G context where we have new actors entering the value chain, new types of devices, etc. It is likely that the findings of Task 2.2 on Trust Model will require a rather different architecture. For example, we know for a fact that the original trust model in the inter-operator networks (designed for a small number of large national operators) is already now problematic, e.g. impersonation and exfiltration of sensitive information on SS7 networks. Even *if* Task 2.2 produces a trust model 100% compatible with the (implicit) model in 3G/4G, then we have at least obtained better evidence of the soundness of the architecture.

Virtualization: The domains defined in [ts23.101] are defined by "physical grouping". A 5G security architecture must have stronger focus to "logical groupings" due to e.g. virtualization.

Management: Management is left completely outside the scope of [ts33.401]. Secure orchestration and management of virtualization, general forms of security management (e.g. monitoring) are absolutely fundamental for 5G to operate robustly. Without it, we cannot achieve the robustness required by telecom regulatory constraints in the presence of general ICT and cyber threats.

Cyber threats: If mission critical services (health, transport, industrial automation, etc.) are going to use 5G networks, this could also act as a "magnet" for more general cyber threats. The damage done (even loss of life) goes beyond the worst imaginable impact on the "mobile broadband" type services that we see today.

1.3 5G-ENSURE Methodology

We have already touched upon the dependency between this work (Task 2.4) and the work with trust models (Task 2.2) and Risk-analysis and requirements capture (Task 2.3) and how we address these in a two-phase iterative approach. Backing up one step further, one can also make the observation that all the tasks 2.2, 2.3, and 2.4 depend on the use-case study performed by Task 2.1. There is no iterative update of the use-cases during the 5G-ENSURE project. It is therefore relevant to ask if there is a risk of failure to capture certain use-cases that might have led to a slightly different trust model and/or slightly different risks/requirements, which in turn would have have led to a slightly different architecture. Of course, such risk cannot be denied. However, the 5G-ENSURE project is confident in that such risks are minimal, motivated as follows.

First, while Task 2.1 has "only" produced 31 use-cases, it should be noted that there is, as such, no "safety in numbers": it is not the quantity of use-cases that is critical, it is their quality. The use-cases have been selected because they seem to "span" the 5G Security problem space well, just as a small number of vectors can still span an "exponentially large" vector space. Furthermore, there has been no censorship: the 31 use-cases comprise *all* use cases suggested by the 5G-ENSURE partners. The expertise available in the 5G-ENSURE consortium makes it highly unlikely that an important corner of the security problem space is missing.

Secondly, the 5G-ENSURE use-cases do not constitute the only background material. For example, the 3GPP "SMARTER" study [tr22.891] comprises over 70 uses cases and the overall 5G-PPP consortium has produced hundreds of use cases. A special use-case harmonization activity has taken place in the 5G-PPP with

representation from 5G-ENSURE. The conclusion so far is that no gap has been identified. If new use cases are brought wirwrad they will be cconsiderd in the remaing work.

1.4 Abbreviations

3GPP	3rd Generation Partnership Project
3P	3rd party
5G-PPP	5G Infrastructure Public Private Partnership
AAA	Authentication, authorization and accounting
AN	Access Network
APT	Advanced Persistent Threat
AS	Access stratum
BEREC	Body of European Regulators for Electronic Communications
CN	Core Network
DoS	Denial-of-service
DDoS	Distributed Denial-of-service
HAPS	High Altitude Platform Stations
HN	Home Network
IoT	Internet-of-things
IP	Infrastructure Provider
ITU	International Telecommunication Union
LTE	Long-term Evolution
ME	Mobile Equipment
MEC	Mobile-edge computing
MNO	Mobile Network Operator
MT	Mobile terminal
MTC	Machine-type Communication
NAS	Non-access stratum
NFV	Network Function Virtualization
QoS	Quality-of-service
RAN	Radio access network
RAT	Radio access technology
SDN	Software defined networking
SDR	Software defined radio
SLA	Servuice Level Agreement
TE	Terminal equipment

TN	Transit network
UE	User equipment
UICC	Universal integrated circuit card
USIM	Universal subscriber identity module
VMNO	Virtual mobile network operator
VNF	Virtual network function

2 The Security Landscape of 5G

5G is being developed with new concepts and capabilities to enable new business models for mobile operators to provide enhanced applications and services to mobile network subscribers. In order to ensure that 5G fulfils its promise, all security matters accompanying the 5G architecture need to be addressed. This section intends to highlight the security characteristics of 5G capabilities; 5G's business and trust model; high-level threats to 5G systems; 5G security problem definition; and the regulatory aspects of 5G.

2.1 Security Characteristics of 5G Capabilities

5G will embrace new technologies and concepts as a significant evolution from 4G. It will employ virtualisation technologies to enhance its infrastructure's flexibility and scalability. Thanks to SDN, it is also refining its network slicing concept to more dynamically offer various kinds of services to VMNOs and users. Besides, 5G aims to enhance its support for low-latency, location-aware, and network-aware applications with mobile edge computing, as well as its support for various kinds of IoT applications. Finally, 5G will to higher extent incorporate the support of other non-3GPP network access technologies to enlarge its service reach to various UEs.

All these new capabilities are exciting and will bring significant benefits to operators, application services and end users. However, they also bring new security challenges and characteristics that needs to be addressed to ensure a successful deployment of these capabilities. This section will discuss the security characteristics of the 5G capabilities.

2.1.1 Virtualisation

5G intends to leverage virtualisation to be cost-effective in infrastructure deployment, flexible in scaling, and dynamic in providing new services. Virtualisation is the underlying enabling technology for network function virtualisation, network slicing, and mobile edge computing in 5G. The current security support for virtualisation, developed and deployed in public and private clouds, serves as the fundamental protection of 5G's virtualised infrastructures, functions and services. However, as a mobile network has a more stringent security requirement than a typical cloud provider in terms of service availability and data privacy, 5G observes in particular the following security characteristics of virtualisation.

- Integrity of virtualisation platform: The integrity of virtualisation platform is the root of trust for virtualised functions and services in 5G.
- Authentication of software entities: It is important to authenticate virtualised network functions, mobile edge applications, and other software entities running in the virtualisation platform to protect against attacks by impersonating software entities.

- Isolation of resources: Runtime memory, data, I/O and other assets of each software process needs to be isolated from the others to ensure that they are not leaked, misused, or corrupted.
- Compliance monitoring of resources: While technical monitoring of resources is self-evident, there could be compliance requirements that the virtualised resources also have to meet. These could, for instance, relate to the geographical locations of the resources and whether such locations are acceptable for the service in question. One might also need to get assurance that certain level of security is offered, instead of resorting to a configuration with weaker security requirements. Some components, such as VNFs, might also need to be certified in order to get the assurance of their proper functions.

2.1.2 Network Slicing

Network slicing (and further sub-slicing) could be created to use some portions of the underlying network to provide network services with particular properties.

Slicing can be utilized in such complicated cases where more than one Virtual Mobile Network Operator (VMNO) share the same 5G physical network which is operated by a provider of virtualized infrastructure. The VMNO's may control their own slices while they provide sub-slices to their customers. The main security characteristics for network slicing are listed hereunder.

- Micro-segmentation to control and prevent anomalies: Anomalous behaviour in an SDN network can be easier to detect and to respond, if the 5G network system is (virtually) divided into smaller parts, i.e., network slices, sub-slices and micro-segments. Through this approach the surface for attacks and threats can be reduced significantly.
- Extremely secure services: Micro-segmentation could provide an even more fine-grained approach than traditional network slicing and with micro-segmentation it can be possible to create extremely secure segments where more granular access controls and stricter security policies can be enforced.
- Trustworthiness of sliced system: For advancing security of the sliced system, the authenticity and integrity of the received data and commands in each slice must be ensured. Furthermore, to control the access between slices, security mechanisms must be able to check, if the received data/commands, originated from within the slice or not (from a legitimate entity). In other words, it must be able to check its trustworthiness, to prevent access from other slices.
- Satisfying SLA objectives: The security system must also ensure that the different SLA objectives for the different slices are met. The SLA objectives will be different depending on the slice's use case (e.g. autonomous driving, health applications, massive IoT, real-time 4K video broadcasting, etc.). To control the varying and sometimes conflicting SLA objectives, the SDN controller should support policy-checking functions.
- Slice isolation: In a 5G network, the isolation of slices (isolation assurance within 5G nodes) must be ensured. This assurance must be provided at two levels, at security level (threats propagating through the slices) and at resiliency level (faults in the physical infrastructure propagating through the slices).
- Physical isolation: A compromised slice may compromise the security of other slices sharing the same physical 5G nodes.
- Limited physical resources: Unavailability of a physical network resource (physical 5G node) serving the slices, due to intentional or accidental disruption, may propagate to unavailability of those slices (a.k.a cascade effect).

- Data integrity: Integrity and authenticity of the data/commands uploaded/downloaded by a 5G controller/object must be ensured to avoid any security issues.
- Resiliency through real-time monitoring and controlling: Resiliency of a sliced system should be ensured to prevent cascade effects between different slices. This can be done by checking in real time which part of the physical infrastructure is ensuring the integrity of a given slice topology and by proposing migrations upon detection of vulnerable, attacked, compromised or affected physical resources. For that, it is necessary to support on-the-fly retrieval of the dynamic dependencies between the slices and the physical infrastructure in order to assess the propagation of faults and attacks in a given slice.

2.1.3 Mobile Edge Computing

Mobile edge computing (MEC) is a capability of hosting third-party applications at the mobile edge on mobile edge hosts deployed at radio nodes, aggregation points, or the edge of the core network. MEC creates new opportunities for 5G operators and applications. First, it allows better support of low-latency applications by placing them in close proximity of their users, avoiding having application traffic traverse the core network. Second, mobile operators can provide mobile edge services, such as location and radio information, to third-party mobile edge applications so that they can optimise their performance and responsiveness. The security characteristics of MEC can be observed in the following aspects.

- Mobile edge hosts outside the mobile operator's premises: Since the data traffic of mobile edge applications do not go through the core network, it seems that lawful interception and traffic accounting need to be performed at the mobile edge hosts. If the mobile edge hosts are deployed at radio nodes or aggregation points, lawful interception and traffic accounting would very likely have to be conducted outside the mobile operator's premises, in a stadium, in a shopping mall, on a campus, on a hill, on a rooftop, etc. This increases the risks of exposing the lawful interception and traffic accounting functions to attacks for illegal eavesdropping, fraudulent billing, etc. Besides, with numerous cloudlets (mobile edge hosts) outside the operator's premises, the physical protection of MEC's virtualisation environment is more challenging than a standard cloud environment.
- Provision of mobile edge services: In MEC, the mobile operator can provide mobile edge services to third-party mobile edge applications. It is necessary that the mobile edge services are capable of authenticating the mobile edge applications to ensure legitimate access to their services. This may in turn pose threats to credentials exposed outside the core network.
- Service continuity of mobile edge applications upon UE handover: In the event of UE handover, to maintain the low-latency communication and context awareness of the mobile edge application, it is necessary that the UE be served by the closest instance of the mobile edge application. Thus service continuity has to be supported either by an application context transfer from the current instance of the mobile edge application to the next instance, or by the transfer of the mobile application instance itself from the current mobile edge host to the next one. In either case, it is necessary to assure that the source and the destination mutually authenticate each other, and that the transfer is secure.
- UE authentication and re-authentication: Low-latency communication is one of MEC's key promises. The current UE authentication and re-authentication approach in pre-5G mobile networks may need to be enhanced to minimise or eliminate their disruption to low-latency communication, in particular during a UE handover.

2.1.4 Internet of Things

One of 5G's key objectives is to support Internet of Things use cases. IoT use cases often involve a potentially large number of IoT devices, from a few home network devices to hundreds of thousands of smart meters. For 5G, IoT presents the following security characteristics.

- Surge of network signalling: In many IoT use cases, the IoT devices behave similarly and very often act simultaneously. When such IoT devices are numerous, this could pose a security challenge to 5G because the 5G network must deal with sudden surges of network signalling and application traffic. The network needs to gracefully sustain the overload without breaking down.
- Distributed denial of services (DDoS): IoT is becoming a new attack vector for DDoS. Whether the attack target is the mobile network or an application, the mobile network will face a surge of network signalling and application traffic. As indicated above, such behaviour is very typical of IoT use cases. Even non-malicious device malfunctions could result in attacks. Thus it is important that 5G can distinguish DDoS from normal IoT behaviours to protect the network resources.
- Extremely constrained devices in a network with strong security algorithms: IoT devices are often designed to operate with extremely low power. Therefore, their processing power and memory size can be limited and they may not be able to support strong security algorithms. Even their radio interface can be non-3GPP, supporting only ZigBee, Bluetooth, or WiFi. Such constrained IoT devices may not be able to access 5G networks directly themselves, and in that case special 5G user equipment may act as IoT gateways for groups of IoT devices. With this approach, each IoT device may still establish itself a point of presence and unique identity in the 5G network and enable itself a service differentiation (such as specific QoS). 3GPP specifies MTC-IWF and Service Capability Server (SCS) to support MTC.
- Group-based authentication and security: In many cases IoT devices may advantageously be treated as groups based on their physical location, type of sensors or actuators, type of application, or other factors. Such device groups could perform simultaneous authentication through an IoT gateway or a mobile device which acts as a relay. This approach could strongly reduce the AAA overheads since each device need not execute the complete protocol individually. In group communication setting also arises the need for securing multi- and broadcast traffic with specific security requirements.
- Impact of high latencies and low access priorities on authentication: For supporting IoT, 3GPP now develops specifications for low access priority, extended access barring and high latency communication. The UEs can be configured with low access priority, which means that much longer delays are tolerated when accessing the network, while high latency communication allows mobile-terminated communication with UEs running in power saving mode. These changes may require enhancement of current authentication procedures.

2.1.5 Heterogeneous Network Access

5G is expected to bring more flexibility to network access. Not only several different kinds of technologies could be used to provide the radio access, but also authentication would be more dynamic and local for access to 5G services.

- Enhanced identity protection: While Internet of Things is one major driving force for the convoluted radio access, satellite is also likely to play part in providing heterogeneous access, especially in sparsely populated or hard-to-reach locations. A challenge is to ensure that the same level of protection is provided in terms of authenticity and identity protection to these complementary technologies. Pseudo-identifiers could be used to provide better anonymity against tracking of clear

text device identifiers, but more advanced key management solutions with perfect forward secrecy could also be used to mitigate any effects of key leakage.

- Identity services interoperability: When it comes to identifying the subscriber identities, there is likely to be more interactions with different kinds of identity providers. MNO could interact with the AAA services of an enterprise or a satellite provider to authenticate the users, while it could also provide identity management services to third parties.
- Dynamic roaming: Dynamic roaming would allow access even in the cases where static roaming agreements do not exist between the home operator and the serving network. In other words, a subscriber might have service needs which could not be covered by the current agreements of the home operator. Such use cases could involve, for example, satellite networks to provide resiliency to improve the decreased level of service due to disruptive environmental conditions such as earthquakes, or capacity overload such as temporary surge of number of users. This calls for mechanisms to establish sufficient trust and assurance of compensation between the network entities.
- Strong accountability: In order to address spoofing concerns, there should be a stronger linkage between the use and the identities. For instance, in the above mentioned dynamic roaming use case, the home operator could get assurance that a subscriber trying to get access to the network is genuine (the same non-repudiation assertion regarding compensation could be then given to the visited network as well). One way of achieving this could be through cryptographic identifiers. Such identifiers also make it easier to bind user related signalling messages to the user, so that it is harder to try to spoof user identity in order to incur costs which are not legitimate. The introduction of cryptographic identifiers and digital signing of signalling could also be used to mitigate denial of concerns in cases due to spoofed signalling messages.

2.1.6 Critical Service Support

Since 5G targets applications related to industrial automation, public safety, vehicle-to-vehicle, and communication with cyber-physical systems, the needs in the area of robustness, resilience high availability and fault-tolerance will be more profound, ranging all the way from radio-access to back-bone transport.

2.2 Business and Trust Model Summary

Based on the reality of 4G network actors and what is planned for 5G (described both above and in the use case analysis below), the following list of 5G actors is reproduced from [d2.2]. New actors compared to the 4G setting are highlighted with a "[5G]" label.

- **Network equipment manufacturer**
  - Terrestrial equipment manufacturer
  - [5G] Satellite equipment manufacturer
- **Infrastructure Provider**
  - [5G] Virtual infrastructure provider (VIP), providing infrastructure as a service (IaaS)
  - [5G] Satellite/ High Altitude Platform Stations (HAPS) provider
- **Network software provider;** commonly also the network equipment manufacturer
  - [5G] Virtual network function (VNF) provider
- **Interconnect network provider** (provides a network linking one network operator to another)

- **Mobile Network Operator** (MNO) (taking the role of "home" or "serving" operator); commonly also the infrastructure provider
    - Virtual mobile network operator (VMNO) who purchases bulk capacity from MNOs and may (or may not) have their own HSS
    - [5G] Virtual mobile network operator (VMNO) who purchases SDN slices from an Infrastructure Provider
    - [5G] Factory or enterprise owner operating a AAA in a network linked to a (V)MNO
- [5G] **Satellite Network Operator**; commonly also the satellite/HAPS provider
- [5G] **Network access provider** (uses the services from one or more Satellite/Mobile Network Operators to provide bulk transmission resources to Service Providers)
- **Service Provider**; commonly also the (V)MNO
    - [5G] Over-the-Top (OTT) service provider
- **User equipment manufacturer**
    - Phone manufacturer
    - USIM manufacturer
    - [5G] Sensor manufacturer
    - [5G] Robot manufacturer
- **User equipment software developer/provider**
    - User equipment operating system developer/provider
    - User equipment application developer
    - Application store provider
- **End user**
    - Common phone users (Service Provider subscriber)
    - [5G] Wireless Sensor Network (WSN) owner/operator
    - [5G] Employee of enterprise
- **Regulators**, law enforcement agencies

The precise relationships between these actors will need to be defined and clarified as the complete 5G architecture is determined. Potentially, new actors ill also be added. This is discussed in more detail and some initial examples provided in Section 6.

2.3 High Level Threats to 5G Systems

In this section, we present high level threats to architectural components of 5G systems. Different risks and their assessment is discussed in Deliverable 2.3 [d2.3] and risks are derived from the analysis of the subset of 5G-ENSURE use cases.

The deliverable [d2.3] discusses several threats derived from the analysis of the subset of 5G-ENSURE use-cases. In this section we only summarize high level threats to 5G systems from architectural perspective. In particular, we outline threats to different entities and interfaces of the 5G security architecture.

2.3.1 Smart Jamming Attacks

The current 2G, 3G, and 4G architectures address persistent jamming attacks in the sense that they auto-recover when attacks stop. Their design also aims to prevent that non-persistent attacks have persistent side-effects, see [LTE_book]. Recent research work has however demonstrated such types of effects in 4G

networks [altaf]. Due to requirements of next generation emergency response communication infrastructure and potential cyber security threats, the 5G system needs to consider protection features against smart jamming attacks. One example of such threat from [roger], local DoS attack via smart jamming against 5G services around national stock exchange or large corporation's headquarters. Although such kind of Advanced Persistent Threats (APT) may be detected and unrealistic, however, due to reliance on mobile communication architecture for public safety is an important factor to consider such threats during the security design phase.

2.3.2 Radio Access Network Threats

Security of the subscriber data and availability of radio networks is of utmost importance in the communication. With 5G networks, there is expected to be a massively increase in the number of connected end-devices. Furthermore, new requirements of 5G business use cases would play a critical role in influencing the design of access network security. For example, some services require low latency or high availability. Similar to earlier generations, trade-offs between availability and performance will play a critical role in addressing threats arising from signalling overload (denial of service) at the radio access network. Such primary threats are denial of service attacks originating from very large set of connected or compromised IoT devices, transfer of radio encryption keys within the networks, and attacks against user plane integrity [ngmn2]. Furthermore, threats to anonymity and privacy of data transmitted between the end devices and radio access network require adequate protection features.

2.3.3 SDN/NFV Related Threats

Security is a key issue in realizing business use cases of SDN and NFV technologies in 5G systems. We summarize threats and impact assessed by the ENISA [enisa] and NGMN working group [ngmn1] in the following two subsections:

2.3.3.1 SDN threats

The SDN threats can be categorized into three types – Application Plane threats, User Data Plane threats, and Control Plane threats. In Application Plane, threats impacting the network services are information leakage, tenant impersonation, communication hijacking, and API abuse. The Control Plane threats affect centralized management of the network policy by network manipulation threats. User Data Plane threats primarily consist of DoS attacks and communication hijacking. The recent vulnerabilities [cve2, ruxcon] against SDN infrastructure indicate a need of careful attention of minimizing impact of such threats on overall 5G communications network.

2.3.3.2 NFV threats

The NFV threats include abusing communication function between network slices and their components that are required for signalling and management, impersonation attacks against network slice manager, denial of service attacks by exhausting network resources, and information leakage attacks.

2.3.4 Location Based Attacks

Location based services are expected to be heavily utilized in 5G networks. Leakage of privacy sensitive information related to the subscriber from such services pose a noteworthy threat, for example, if the device or subscriber identities can be intercepted when not protected during services like LTE Direct or proximity services. Further, 5G radio network features such as 'always ON' positioning [always-on] is beneficial for supporting Internet of Things or autonomous driverless car navigation. However, privacy and security threats to such location based 5G enhanced services must be addressed during the design phase. Another example

of location enhanced 5G service is smart drones delivering medicine supply by exploiting support of 5G networks. In such cases, adequate protection to location based 5G services is needed to realize business needs.

2.3.5 Small Cells Security

Ultra-dense deployment of small cells is one of the key aspects of 5G system [Speed-5G], but they also introduce security threats. Recent studies [Borgaonkar, blackhat] discussed realistic threats of small cells against the current 3G and 4G network infrastructure. These attacks indicate issues in small cell security architecture, and trade-off between cost and security. Combination of Low cost hardware and software security issues in small cells may pose significant threats to 5G network.

2.3.6 Cyber Security Threats

The security architectures of current 2G, 3G, and 4G networks do not address threats to the architectural end-nodes. Some examples of such end-nodes would be software and hardware packages running on smartphones or small cells. Recent incidents and security issues [android, huawei1, huawei2, p1] indicates emerging cyber security threats affecting main core security principles of mobile networks. We summarize such cyber security threats to few important architectural nodes. Although all these cyber security threats are not realistic to fully address, a careful consideration of their impact on 5G networks can be beneficial during the security design phase.

2.3.6.1 Malware attacks on 5G devices

Nokia reported that smartphone infections in the first half of 2016 are accounted for 78 percent of the infections detected in the mobile network [nokia]. Typically, a malicious application installed on the smartphone could steal private information stored on the 5G devices. Such types of information stealing threats are out of scope from the architecture point of view. However, these infected 5G smartphones may impose a significant threat to the mobile network by generating signalling overhead (denial of service attacks from mobile botnets). Also other compromised 5G IoT devices running similar smartphone OS can increase a threat level. Moreover, recent research indicates that baseband vulnerabilities can be exploited to compromise smartphones to perform man-in-the-middle attacks [golde].

2.3.6.2 Malware attacks on AAA infrastructure

The AAA infrastructure handling subscriber authentication data including keys is an essential asset to be secured. Recent incidents highlighted possibilities of compromising a vendor's authentication and encryption keys infrastructure by a sophisticated malware [gemalto]. Such malware affects authentication and confidentiality aspects of mobile networks.

2.3.6.3 Malware attacks on SDN infrastructure

The SDN applications deployed by the mobile network operator can compromise the SDN controller which is responsible for managing the operational network. These malicious SDN applications can cause denial of service attacks or enable privilege access to the network state or turn into a botnet [sdn1, sdn2].

2.4 Regulatory Aspects

Regulation is not a new aspect for mobile networks. 5G network, like the previous mobile network generations, must therefore undergo a set of regulations including new and old ones. Its implementation can

be different. In this section we first discuss the spectrum regulations. Then, we highlight the 5G ecosystem environment in order to show that a responsibility re-allocation is needed and a new perception of neutrality is required. We also consider the roaming example to demonstrate the importance of responsibility re-allocation and liability chains. Finally, we provide the main Lawful Interception requirements namely the confidentiality of users' data and services.

The new 5G use cases like broadband access in dense area or Massive IoT, are expected to dramatically increase exchanged data volumes. This is one of the main reasons to have new spectrum for 5G. However, the frequency bands are very densely used and free spaces are rare. In the US and Europe a major source could be the spectrum released by the migration from analogue to digital television. The frequency bands are designated by the ITU-R or by individual regulatory bodies. For 5G, the choices include new spectrum below 6GHz, as well as spectrum in higher frequency bands [bt], [ericsson]. The choice of spectrum either for licensed or unlicensed use is important because it affects the cost of equipment, hence the price of services, coverage, and inter-operability.

The 5G is assumed to be the future Internet. Unlike the previous mobile network generations, so called 4G/LTE, that provides homogeneous connectivity to customers, 5G is expected to be versatile: it will encompass various access network technologies, i.e., fixed access, radio access (3GPP RANs and Non-3GPP RANs), and provide connectivity to heterogeneous services such as mass market, IoT and Public Safety. In this context, the telecom Industry warns, in what they called "5G Manifesto", that the current Net Neutrality guidelines, as put forward by BEREC, "create significant uncertainties around 5G return on investment, concurs with Industry verticals that the implementation of Net Neutrality Laws should allow for both innovative specialised services required by industrial applications and the Internet Access quality expected by all consumers", and points out "the danger of restrictive Net Neutrality rules". In addition, they consider the new "concept of 'Network Slicing' to accommodate a wide-variety of industry verticals' business models on a common platform, at scale and with services guarantees" [bt]. In addition, 5G would imply new technologies, like virtualization, SDN and NFV. Thus, a new 5G ecosystem and new actors / roles / entities are emerging. In addition to traditional roles, namely Mobile Network Operators (MNO), Service Providers (SP) and end users, we may have new roles such as Network Infrastructure Provider, Virtualized Infrastructure Owner, Virtualized Function Provider or Slice Owner. All those roles can be for instance assured by the MNO. In this case, we end up in the same case as previous mobile network generations where the MNO is the main owner of the mobile network and also main responsible for any issue (financially responsible). In the other case, where the roles are allocated to different entities/actors, main-owner-responsible system (applied in the previous mobile network generations) is not any more relevant. Consider for instance, a VNF running within the network. This VNF is provided by a VNF Provider A, runs within a slice that belongs to a Slice Owner B and managed by a Slice Provider C. This VNF is managed by a VNF Manager D and runs over hardware from Infrastructure Provider E. If this VNF has a security or functional issue (e.g., unavailability, underperformance, security breaches, isolation breaches, non-compliance with security policies), is the responsible A or B or C or D or E or, A and B, or etc. How can we designate the responsible? In such a complex environment, it is important to set a traceability system that enables to identify the source of an issue and whom responsibility. Thus, liability chains are important. This system can be based on remote or local attestation of a given property like the code integrity, the trustworthiness of the execution environment, the isolation of a slice and so forth. Certifications can also be a tool to build liability chains. Naturally, in addition to the technical solutions, agreements between different actors are needed.

Now consider the example of roaming, in the previous mobile network generations, we distinguish domestic and international roaming. Domestic roaming is used between national operators to offer a better coverage

to end users in the same country. International roaming is used between mobile network operators from different countries. Whatever is roaming type, the roaming operator must rely on the choices made by the 'visited' operator running the network in that area. In a 5G context, roaming takes new dimensions. We can have roaming agreements at different level, e.g., between Slice Owners, between Slice Providers, between Infrastructure Providers, etc. For instance, we can have a roaming agreement between two Slice Owners even if they are running over the same infrastructure.

Whatever is the ecosystem, lawful interception requirements remain the same as in Section 13 of [d2.1]. Indeed, the LI dilemma is ensuring the end user and services privacy (namely the confidentiality) vice versa the ability to answer any LI requirement. This mainly results in the following elements. The mobile network operator must ensure that only those under surveillance are wiretapped, e.g., authorities cannot wiretap users/entities not on the list. Only the mobile network operator must be able to trigger a Lawful interception. These imply a strong isolation requirement in the mobile network to prevent fraudulent network access and abusive use of resources. In addition, only concerned entities (i.e., the mobile network operator LI service and Law Enforcement Agency) have access to the list of the wiretapped and collected data. The mobile network operator must be able to answer any LI request without requiring any third party. This operation must not be detectable through observation or quality of service. Finally, in case of an end-to-end encryption managed by the network, the mobile network operator should be able to deliver plain data or the encrypted data along with the decryption key.

2.5 The 5G Security Problem Definition

Here we aim to capture the core essence of 5G security as problem statement, summarizing the security landscape and its threat situation as laid out above. Obviously, when facing the new 5G security challenges we must not neglect to provide at least the same level of security that users have been able to obtaining in existing 2G-4G systems. This leads us to the following problem statement.

> *Maintain 4G success as a trustworthy mobile broadband service, and extend the security functions against a landscape of advanced cyber-security threats evolving from new use cases, such as critical communication services and IoT. Create highly scalable, flexible and efficient security infrastructure and security protocol design that fulfils the security needs of various stakeholders of the 5G ecosystem, including the needs of the traditional telecom stakeholders such as subscribers, network operators, and regulators, as well as the needs of emerging new stakeholder such as network owners from private and public sectors, virtual mobile operators, and telecom-cloud infrastructure providers. Allow this multitude of 5G user categories to securely share common, virtualized network infrastructures.*
>
> *Build solid security for the new 5G radio technology, the new 5G access networks composed by diverse set of different access technologies, the virtualized 5G core networks, and the global network-of-networks composed of private, public and virtual network components. Separate the security of access and core networks[1] allowing future-proof independent evolution of new radio, access network and core network technologies.*

[1] This does not imply breaking "end-to-end principles" since it only affects lower layers. End to end security at e.g. IP or transport level is still possible.

Pay special attention to present and new types of end-user equipment such as sensors, actuators, groups of small devices, relay-nodes, smart phones, or any other end-user equipment, and make unauthorized tracking, interception or any other violation of privacy of the subscribers infeasible.

We will in the next section break this down into more fine-grained *security objectives*. The reader familiar with the Common Criteria assurance standard [cc] may recognize this approach, though we here do not use the full formalism. For example, some of the objectives below (e.g. Section 3.8) would in the CC case probably be expressed as policies.

3 5G Security Objectives

The derived 5G security objectives fall into the following categories.

3.1 General

O1.1 Where possible, 5G security should be decoupled from specific physical deployments, focusing on defence-in-depth, in particular self-protection of assets, limiting dependency on protection at network, site, or node perimeter

O1.2 In a logical or physical part of the network, governed by a specific security policy, further fine-grained security policy enforcement should be possible based on mechanism such as e.g. micro-segmentation

3.2 Security-relation to Legacy Systems

O2.1 5G must provide a security and privacy level higher or at least equal to the security and privacy level in 4G.

O2.2 5G security should not be negatively affected by the security of legacy systems with which it interworks.

O2.3 5G must provide solutions for security and privacy breaches identified in the previous mobile network generations such as the IMSI and IMEI unauthorized tracking or the denial of service provoked by the unsecured mobility messages (i.e., TAU messages).

O2.4 5G must enable seamless interworking of different network technologies, mobile, fixed as well as satellite [bt] without exposing the security level of each of these technologies to new threats.

3.3 RAN/RAT Security

O3.1 5G Radio Access Technologies (RATs) should exhibit higher resistance to advanced jamming and other DoS attacks than current RATs.

O3.2 5G RATs should be able to provide user plane data integrity

O3.3 5G access security should allow for high efficiency in authentication and security set-up, supporting ultra low latency services

3.4 Secure Virtualization

3.4.1 General

O4.1 5G must enable a secure, reliable, and traceable sharing of network resources (i.e., compute, storage and network) between the various services having vastly different requirements such as reliability and low latency for tactile remote surgery and availability for massive IoT services.

O4.2 5G security must be dynamically scalable in order to easily and securely enable the changes required to ensure any new use cases, new trust models and new service delivery approaches.

O4.3 5G infrastructure components should support necessary root-of-trust functionality

3.4.2 Slicing

O5.1 Slices must provide inter-slice isolation of sensitive data, approaching that of physically separated networks

O5.2 Slices must provide strong isolation of (virtual) resources allocated to different slices

O5.3 Slices must support configurable security and be able to provide slice-unique security services as required by specific services and applications

3.5 IoT Security

O6.1 5G security protocols must scale to support massive IoT and must not have negative security impact on non-IoT services

O6.2 5G security must support also protocols efficient enough for resource constrained devices

O6.3 5G connected devices must be able to adequately protect critical data such as subscriber credentials against threats of unauthorized access and/or modification

3.6 Security Management

O7.1 5G must support traditional UICC/USIM based management of subscriber credentials

O7.2 Management and storage of alternative credentials by an external vertical AAA must be done with a security level commensurate with both that of the vertical application as well as the operator business partner

O7.3 5G systems must support security monitoring capable of detecting advanced cyber security threats and support coordinated monitoring between different domains and systems (e.g. mobile and satellite)

O7.4 5G systems and components must provide strong mutual authentication and authorization

O7.5 5G systems and components must provide functionality to mutually assess the trustworthiness before, and during interactions

O7.6 5G systems' interactions must be auditable and produce evidence of liabilities

3.7 New Business and Use-cases

O8.1 5G must be able to deliver and maintain SLA to verticals in terms of: availability, security, latency, bandwidth, access control from an end to end perspective.

O8.2 5G systems must allow secure interworking with external systems, e.g. AAA provided by a vertical or external management systems

3.8 Regulatory Aspects

O9.1 5G systems must comply with regulatory aspects, including those related to Lawful Intercept, user privacy, and customer notification of security breaches.

O9.2 If required by local regulation, 5G infrastructure operator must have means to demonstrate their provided level of security.

O9.3 All 5G security features must be compatible with local regulation.

4 5G Security Architecture Overview

This section defines and explains the main building blocks of the 5G-ENSURE draft security architecture. The core of our proposed security architecture for 5G networks extends and revises the domain and stratum concepts from 3GPP TS23.101 [ts23.101]. These concepts provide different viewpoints of a network and model different network aspects. Before providing in the Sections 4.2, 4.3, and 4.4 the details to both these concepts and their corresponding viewpoints, we elaborate in Section 4.1 on our rationale of choosing TS23.101 as a basis for the 5G-ENSURE security architecture.

4.1 Rationale

TS23.101 identifies and names the reference points and functional groupings appearing of the general UMTS architecture at a high level, from both physical and functional viewpoints. Namely, the physical aspects are modelled using the *domain* concept and the functional aspects are modelled using the *stratum* concept. These viewpoints and concepts proved very useful and are widely used. For instance, the 3GPP system and security architecture defined in TS33.401 [ts33.401] builds upon these two viewpoints and concepts. However, we note that TS33.401 focuses on the functional aspects by using the stratum concept and uses less of the domain concept. This is one reason for lacking a solid anchoring in the trust model. TS33.401 also adds a categorization of security mechanisms by introducing so called *security feature groups*, which are loosely connected to the domains. Details are provided in the forthcoming subsections.

Although 5G networks will be in some regards very different from their predecessors, e.g., through the use of virtualization and support for diverse and critical non-telecom oriented services, they will still share similarities and they will reuse and extend concepts that have proved successful and that are widely adopted. Furthermore, 5G networks must provide some backward compatibility and will use existing network infrastructure to some extent, i.e., standard mobile broadband services will continue to be important and mobile operators will continue to have business agreements among themselves, offering users the ability to roam between home and serving networks, using USIM-equipped smartphones, etc. Reusing and building upon the accepted and well-known concepts and terminology in TS23.101 (also TS33.401 and other standards) helps to understand the similarities and differences better. In particular, the advanced features and capabilities of 5G networks are highlighted. Furthermore, it is more likely to quickly reach a common and coherent understanding, which results in consensus and new or revised standards. Finally, it provides us with the opportunity to clarify or correct earlier work by eliminating some of its misunderstandings or shortcomings that we have identified as part of our work.

4.2 Domains

We begin by quoting the domain definition from TS23.101. Furthermore, we briefly recall the domains introduced in TS23.101.

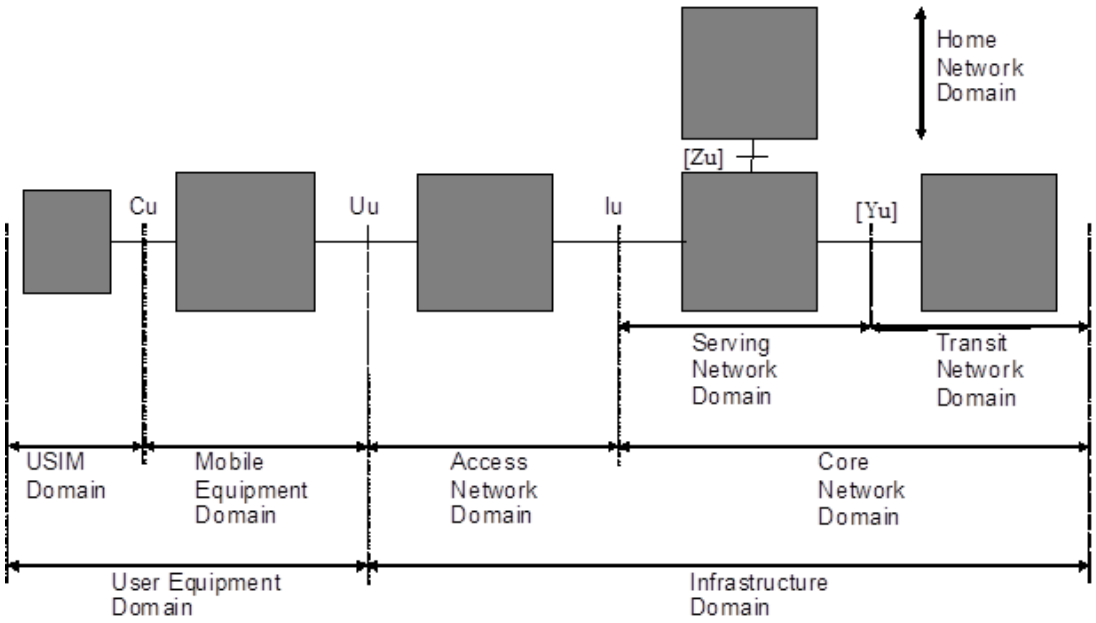

Figure 1: Domains and reference points as defined in TS23.101.

According to TS23.101 (Clause 3.1, p. 5), a *domain* is ”the highest-level group of physical entities”. Furthermore, ”reference points are defined between these domains”. Figure 1, also taken from TS23.101 (Clause 5), provides an overview of the domains and the reference points between them. The following list provides a short description of these domains. Note that domains can have subdomains to provide a more fine-grained grouping of entities.

1. **User Equipment Domain.** User equipment (UE) is the equipment used by the user to access network services.
    a. **Mobile Equipment Domain.** The Mobile Equipment (ME) performs radio transmission and contains applications. The mobile equipment may be further sub-divided into several entities, e.g. the one which performs the radio transmission and related functions, Mobile Termination (MT), and the one which contains the end-to-end application or (e.g. laptop connected to a mobile phone), Terminal Equipment (TE).
    b. **USIM Domain.** The USIM contains data and procedures which unambiguously and securely identify itself. These functions are typically embedded in a standalone smart card. The USIM device is associated to a given user, and as such allows one to identify this user regardless of the used ME.
2. **Infrastructure Domain.**[2] The infrastructure consists of the physical nodes which perform the various functions required to terminate the radio interface and to support the telecommunication services requirements of the users. The infrastructure is a shared resource that provides services to all authorized end users within its coverage area.

[2] This domain name is introduced and used in TS23.101. In recent work however, namely [etsi_nfv], the same domain name is used with a different meaning. We decided to rename this domain of TS23.101 into Network Domain.

a. **Access Network Domain.** This domain consists of the physical entities which manage the resources of the access network and provides the user with a mechanism to access the core network domain.

b. **Core Network Domain.** This domain consists of the physical entities which provide support for the network features and telecommunication services. The support provided includes functionality such as the management of user location information, control of network features and services, the transfer (switching and transmission) mechanisms for signaling and for user generated information.

    i. **Serving Domain.** This domain represents the core network functions that are local to the user's access point and thus their location changes when the user moves. The serving network domain is responsible for routing calls and transport user data/information from source to destination. It has the ability to interact with the home domain to cater for user specific data/services and with the transit domain for non user specific data/services purposes.

    ii. **Transit Network Domain.** This domain is located on the communication path between the serving network domain and the remote party. If, for a given call, the remote party is located inside the same network as the originating UE, then no particular instance of the transit domain is activated.

    iii. **Home Network Domain.** This domain represents the core network functions that are conducted at a permanent location regardless of the location of the user's access point.

The above definition from TS23.101 of the term *domain* is obviously too narrow for 5G networks. In particular, it limits itself to physical network entities and does not account for virtualized network entities, which will play a dominate role in 5G networks. We therefore broaden the scope of a domain as follows.

Definition. A *domain* is a grouping of network entities according to physical or logical aspects that are relevant for 5G networks.

The phrase "relevant for 5G networks" has been added to prevent the introduction of domains covering all sorts of logical aspects, which are however unrelated or only marginally related to 5G and/or networking aspects. A large number of domains would most likely result in a confusing and complicated architecture. Keeping in mind that complexity is usually unfavourable for security considerations, we seek to stop at a small number of domains, which are however already sufficient for capturing the gist of security in 5G networks. Examples of relevant aspects can include type of functionality, trust, (geographical) location, etc.

Furthermore, the respective domains User Equipment Domain, Infrastructure Domain, etc. as defined in TS23.101 also need to be revised and extended. Figure 2 provides an overview of the domains we foresee in 5G networks. It uses the following abbreviations.

Abbreviation	**Meaning**
UE Domain	User Equipment Domain
ME Domain	Mobile Equipment Domain
ME HW Domain	Mobile Equipment Hardware Domain
UICC Domain	Universal Integrated Circuit Card Domain

USIM Domain	Universal Subscriber Identity Module Domain
IM Domain	Identity Module Domain
AN Domain	Access Network Domain
HN Domain	Home Network Domain
SN Domain	Serving Network Domain
CN Domain	Core Network Domain
IP Domain	Infrastructure Provider Domain
TN Domain	Transit Network Domain
3P Domain	3rd Party Domain
IP Service Domain	Internet Protocol Service Domain

In the following, we describe the domains of Figure 2 in more detail.

Note that Figure 2 also depicts reference points between domains (i.e., the green lines). In this draft version of the 5G security architecture we will however not go deeply into these reference points and therefore no names (such as “Cu”, “Uu”, and “Iu” of TS23.101) are yet assigned to these interfaces. Furthermore, Figure 2 contains additional features for domains to account for network slices and trust issues between domains. Slice domains are, for example, depicted as solid and dashed parallelograms inside domains in Figure 2. They are usually transversal to several other domains. We refer to the forthcoming sections for additional details.

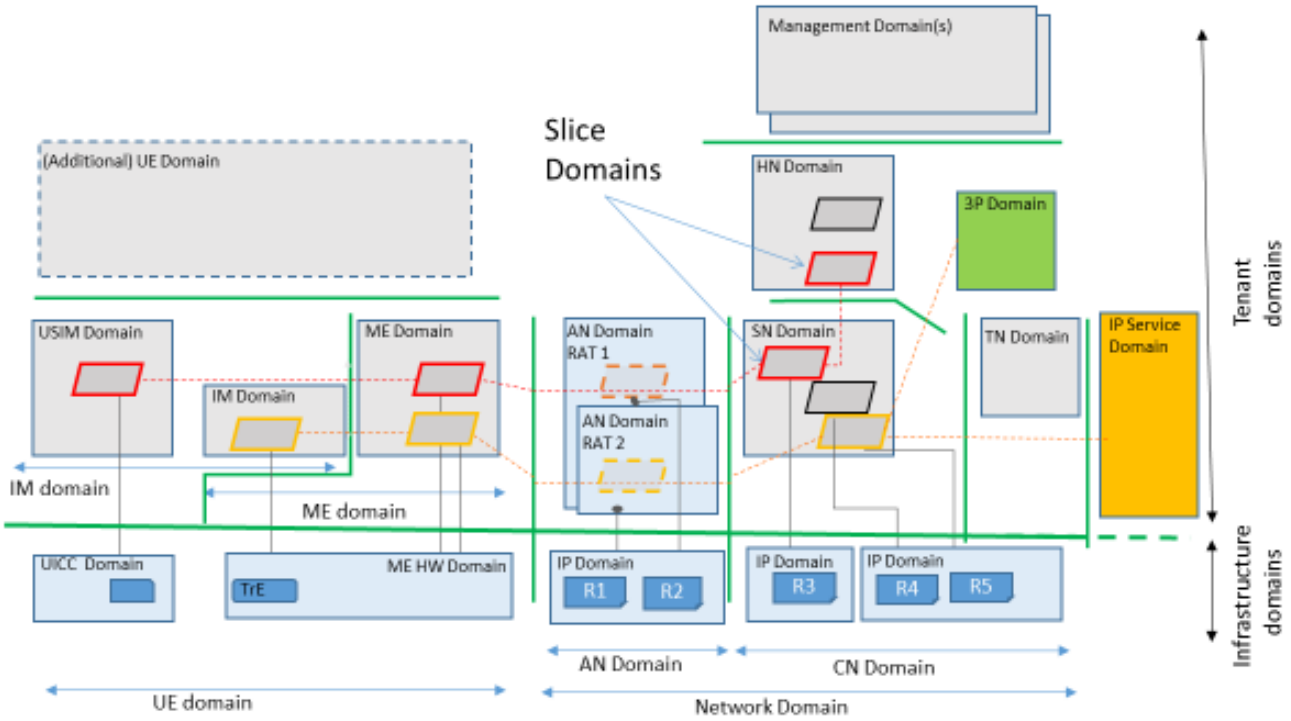

Figure 2: Domains of the draft 5G security architecture. The green lines are used to mark logical or physical communication interfaces between domains, which will be elaborated in a future version of this document. Rombic shapes denotes Slice Domains and lines from slice domains down to Infrastructure Domains denotes that VNFs allocated to a slices make use of certain physical resources.

4.2.1 Domain Types

With reference to Figure 2, we can make a first classification of domains according to whether they are physical or logical, by considering the horizontal green line separating *tenant domains* from *infrastructure domains*. These domains have direct correspondences in the ETSI NFV work [etsi_nfv]. We urge the reader to note that this means that "infrastructure domain" now gets a new meaning, different than the one it had in the previous 3GPP work (as the collection of network-side domains). For this reason, the correspondent to the Infrastructure Domain of TS23.101 is now assigned the new name Network Domain.

We also define the concept of *compound domain*. This is simply a collection of other domains, grouped together according to some 5G relevant aspect, e.g. ownership, joint administration or the like. For the time being, we foresee two new types of compound domains, slice domains and administrative domains, which we will elaborate more in detail below.

4.2.2 Infrastructure Domains

The infrastructure domains focus on the relevant physical network aspects, similar to TS23.101. In other words, the infrastructure domains contain the "hardware" required by both the Network Domain and the UE Domain.

4.2.2.1 UICC Domain

This new domain contains the conventional tamper-resistant module offering protected storage and processing of long-term subscriber credentials in a similar way to the current 3G/4G setting.

4.2.2.2 ME HW Domain

This domain contains the hardware support for the ME (i.e. the MT and TE hardware in the language of TS23.101 such as the radio modem hardware, application CPU, etc.). The ME HW may include trusted execution environments (TEE) supporting other forms of subscriber credentials such as certificates in order to support e.g. factory automation use cases as presented in [d2.1]. In this draft, we do not further elaborate on potential security requirements on such TEE.

4.2.2.3 Infrastructure Provider Domain

The Infrastructure Provider Domain (IP Domain) contains the hardware platforms for the compute, storage, and networking resources required by both the network/telecom functionality and the access (radio) specific hardware. The Infrastructure Provider Domain could be owned or operated by a different entity than the network operator providing its services on top. Note that while Figure 2 has been drawn in full generality, showing a separation between Infrastructure Domains and Tenant Domains in all parts of the network, certain parts, e.g. the Access Network Domain may be less likely to actually outsource the infrastructure to a 3rd party Infrastructure Provider Domain.

4.2.3 Tenant Domains

4.2.3.1 ME Domain

The Mobile Equipment Domain is analogous to the ME Domain of TS23.101, though as mentioned, it now only contains the software parts.

4.2.3.2 USIM Domain

This domain is analogous to the ME domain of TS23.101, though it now only contains the USIM application, not the UICC.

4.2.3.3 AN, SN, HN, and TN Domain

These domains comprise the same functionality as in TS33.401. The only difference is that they now only contain the abstract/logical functionality, provided on top of an Infrastructure Provider Domain (although administratively speaking, the Infrastructure Domain, as well as the overlying AN, SN, HN, and TN Domains may be owned and/or operated by the same entity).

4.2.3.4 IM Domain

The Identity Management Domain is introduced to capture the desired 5G capability to support alternatives to USIM-based authentication, i.e. for industry automation use cases. The IM Domain may contain for example public key certificates. The IM domain preferable obtains security support from a UICC or from a TEE in the ME HW as discussed above.

4.2.3.5 3P Domain

The 3rd Party (3P) Domain is introduced to capture the aforementioned use cases where a third party such as a factory/industry vertical provides its own AAA server and credential management to authenticate M2M devices such as industry robots.

4.2.3.6 IP Service Domain

The IP (Internet Protocol) Service Domain represents operator-external IP networks such the public Internet and/or various corporate networks. Such networks may be partially or fully non-trusted. Note the difference to the more trusted 3P Domain which is governed by a contractual agreement between operator and the third party.

4.2.3.7 Management Domain

Management is largely left outside the previous 3GPP architectures (or assumed “bundled” into the domains themselves), but with more criticality due to virtualization, management of security, etc., it becomes necessary to treat management explicitly, comprehensively, and more consistently. As will be discussed in Section 4.3, we propose to add a new stratum to capture management functions. Is it necessary to also add a management domain? We believe so, since it is becoming more and more common that management is in one way or another outsourced to third parties, e.g., the equipment vendor. In such cases, we need to be able to model trust between the entities and this is easier to capture through a separate domain. Of course, this does not preclude that the management domain belongs to the same administrative domain as the rest of the 5G network.

4.2.4 (Additional) UE Domain

There is already ongoing standardization in the device-to-device communication area, where two UEs are allowed to communicate directly without going through the network infrastructure, except for some initial hand-shake/set up. An additional UE Domain has therefore been added to capture this. The additional UE can of course also be decomposed into domains for the hardware, the UICC, etc., though we chose not to show it in Figure 2.

Note that various group concepts are being discussed in 5G-ENSURE where UEs can be joined into groups. In the present draft version of the architecture, we have chosen not to investigate concepts like “UE group domains”.

4.2.5 Compound Domains

As in TS23.101, and also for other, more 5G-specific reasons, we may also group domains together according to various criteria, thus creating *compound domains* as mentioned above.

4.2.5.1 Legacy compound domains

In direct analogy to TS23.101, we may group domains together according to their "placement" in the architecture, this can be considered as a division in the horizontal dimension of Figure 2.

First, we distinguish between the terminals/devices and the rest of the network by defining, in complete analogy with TS23.101, the *UE Domain* and the *Network Domain*. Note here that our defined Network Domain corresponds directly to the infrastructure domain of TS23.101, the name change necessitated by the conflict with ETSI NFV terminology. Likewise, the Home Network, Serving Network, and Transit Network Domains may be grouped as the *Core Network Domain*. The *Access Network (AN) Domain* is in our case also including different types of accesses, e.g. both WLAN and 5G-radio accesses.

4.2.5.2 Slice Domain

A central feature of 5G is network slicing which we capture by a special form of compound domains.

The *slice domain* enables the network to provide virtual networks, optimized for delivering specific types of services, e.g. an ultra-low latency slice for critical industry automation, a slice optimized for real-time multimedia, etc. Indeed, specific slices could be defined to offer special security services, e.g. allowing special AAA solutions, unified threat management services, strong isolation of information etc. A slice can cover only parts of the network (i.e. part of the CN domain only) but may in other cases be defined end-to-end. It may therefore happen that while the slice is (logically) defined end-to-end, parts of the network are not implementing all the slice-support functions, for example, in a part of the network that uses some legacy equipment. We use the term *slice-aware* to signify if a particular part of the network has full support for slicing. A legacy part of the network, which may thus happen to not be slice-aware could still have functionality relevant for slicing. For example, (legacy) functions related to e.g. QoS could provide useful slicing support, even if other aspects of slicing (e.g. strong data isolation) is not present. In other words, slice related orchestration may still be performed in parts of the network that lacks full slice support. In Figure 2 above we use dotted lines around parts of slices which are located in parts of the network that are not fully adapted to be slice aware. We use a solid line, drawn from the "slice box" to the IP domain, showing that for slice-aware domains, the slice is "anchored" in the domain. See also below.

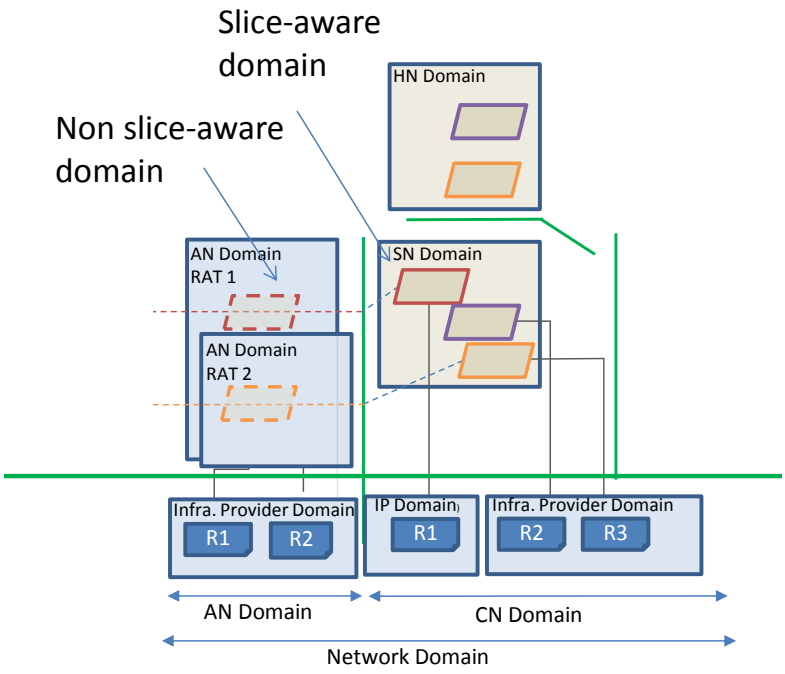

Figure 3: Slice-aware and non-slice aware domains.

Though not shown in Figure 3, it is conceivable that also some Management Domains may contain slices, separating different aspects of management.

4.2.5.3 Administrative Domain

An *administrative domain* is a special case of a compound domain defined not by the presence of a special functionality, but by ownership and/or administration. This is again borrowed from ETSI NFV. While not shown in Figure 2, an administrative domain is simply a collection of other domains, defined by ownership and/or administration. For example, a Serving Network (SN) domain and Access Network (AN) domain may jointly define an administrative domain, e.g. if owned by one and the same mobile network operator.

Note that there is no direct coupling between Administrative Domains and Management Domains. For example, a AN Domain together with an SN Domain may form one Administrative Domain, denoted A, e.g. defined by ownership of one single operator. That operator may however outsource the management of A to a 3rd party (with liability regulated in contract). Thus, management of A may be provided by a Management Domain belonging to a separate Administrative Domain, denoted B.

4.3 Strata

Similar to Section 4.2, we begin by quoting the stratum definition from TS23.101. Furthermore, we briefly recall the strata introduced in TS23.101.

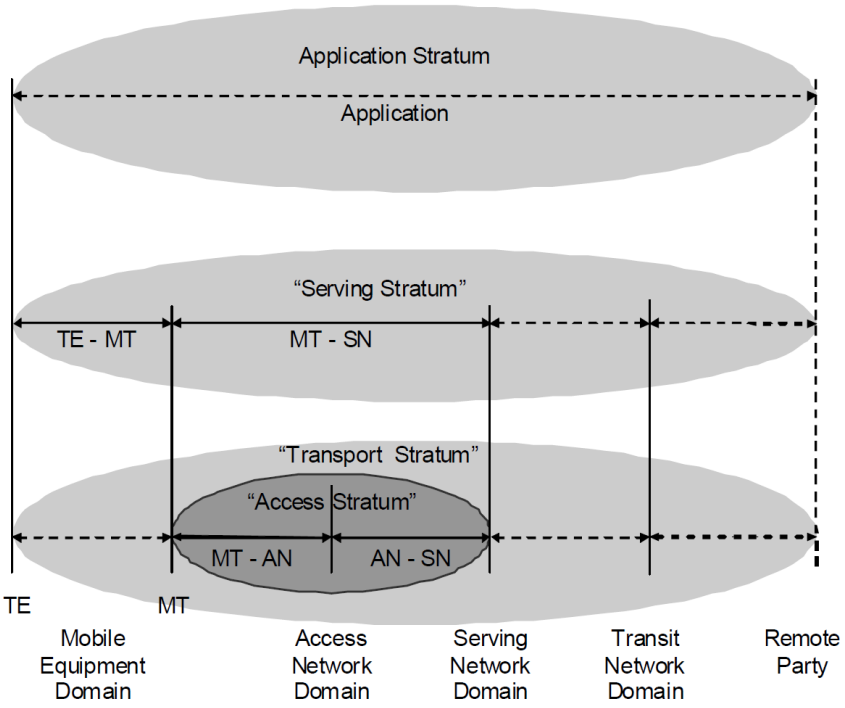

Figure 4: Functional flow between TE, MT, Access Network, Serving Network, Transit Network Domains and the remote party.

According to TS23.101 (Clause 3.1, p. 5), a *stratum* is a "grouping of protocols related to one aspect of the services provided by one or several domains." Figure 4, which is again taken from TS23.101 (Clause 6), shows the following five strata defined for UMTS.. These strata are used in the security architecture descriptions of TS33.102 (3G) and 33.401 (4G).

1. **Application Stratum.** This stratum represents the application process itself, provided to the end-user. It includes end-to-end protocols and functions which make use of services provided by the home, serving and transport strata and infrastructure to support services and/or value added services.
2. **Home Stratum.** This stratum contains the protocols and functions related to the handling and storage of subscription data and possibly home network specific services. It also includes functions to allow domains other than the home network domain to act on behalf of the home network. Functions related to subscription data management, customer care, including billing and charging, mobility management and authentication are located in this stratum.
3. **Serving Stratum.** This stratum consists of protocols and functions to route and transmit data/information, user or network generated, from source to destination.
4. **Transport Stratum.** This stratum supports the transport of user data and network control signaling from other strata through UMTS.
    a. **Access Stratum.** The Transport Stratum comprises the Access Stratum as sub-stratum. This stratum is specific to UMTS and located between the edge node of the serving core network domain and the MT.

Considering the applicability of these strata in 5G networks, one can immediately notice that no stratum is defined for management aspects. In the 5G case, all the (security critical) functionality related to orchestration, virtualization management, security management (monitoring, key distribution, etc.) strongly

speak for the introduction of a management stratum in 5G. We need to be able to model the specific (and usually very stringent) security requirements of the management aspect.

Another issue is that the definition of stratum as "grouping of *protocols* related to one aspect of the services provided by one or several domains", seems to include only communication/signalling aspects through the usage of the word "protocols". Clearly, end-points of a protocol must reside in some stratum (usually the same stratum as that of the "protocol"). Therefore, end-point *functionality* can also be considered to fall inside some stratum. (In fact, that this is the real intention also in TS23.101 can be seen in later discussions in that specification.) Finally, one can note that (big) data, being an important aspect of 5G should be highlighted in the 5G architecture. For example, data communicated within a strata may require protection not only during communincation, but also when stored/processed in the end-points. This leads to the following revised (clarified) definition of stratum.

Definition. A *stratum* is a grouping of protocols, data, and functions related to one aspect of the services provided by one or several domains.

We maintain the same strata as in TS23.101 (albeit with the above extended definition). As mentioned though, we add the Management Stratum.

4.3.1 Management Stratum

This stratum comprises aspects related to conventional network management (configuration, software upgrades, user account management, log collection/analysis) and, in particular, *security management* aspects (security monitoring audit, key and certificate management, etc.). In addition, aspects related to *management of virtualization* and service creation/composition (orchestration, network slice management, isolation and VM management, etc.) belong to this aspect.

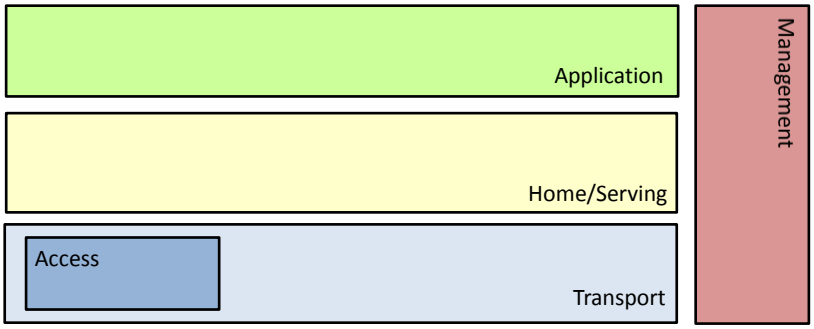

Figure 5: Strata of the 5G Security Architecture.

Figure 5 shows the strata of our draft 5G security architecture, including the management stratum. We have chosen to graphically draw the management stratum vertically, rather than horizontally. That is motivated by the fact that the management stratum will perform management operations on network functions in all of the other strata. For instance, it will comprise protocols like OpenFlow [of] for configuring network components. Obviously, there will also be dedicated protocols, data, and functions related to managing NFVs and network slices.

4.4 Security Feature Groups

TS33.102 and TS33.401 introduces a categorization of security mechanisms by defining so called *security feature groups*. The following five security feature groups are defined.

- **Network access security (I)**: the set of security features that provide users with secure access to services, and which in particular protect against attacks on the (radio) access link.
- **Network domain security (II)**: the set of security features that enable nodes to securely exchange signaling data, user data (between AN and SN and within AN), and protect against attacks on the wireline network.
- **User domain security (III)**: the set of security features that secure access to mobile stations.
- **Application domain security (IV)**: the set of security features that enable applications in the user and in the provider domain to securely exchange messages.
- **Visibility and configurability of security (V)**: the set of features that enables the user to inform himself whether a security feature is in operation or not and whether the use and provision of services should depend on the security feature.

In TS33.102 and TS33.401, it is illustrated in what strata and between which domains the security features belonging to the different security feature groups are present. This is shown in Figure 6 that is borrowed from TS33.401. This is the primary use of the security feature groups in TS33.102 and TS33.401.

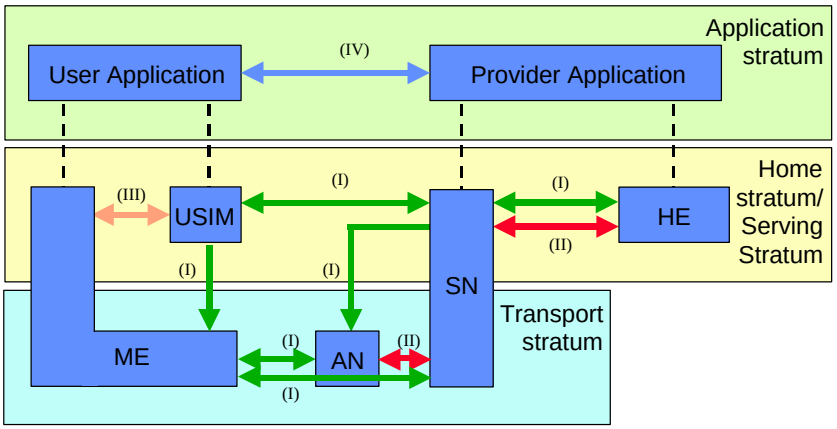

Figure 6: Security feature groups, domains, and strata from TS33.401.

The security feature groups described above are relevant also for 5G but the groups are insufficient and do not fully fit 5G security features such as management, monitoring, and virtualization aspects. We add a security management feature group (VI) to cover also these parts. Furthermore, the visibility of security in group V is focused only on the user, but in 5G this is relevant also for network nodes that needs to know whether security of other entities with which they communicate is properly enabled/configured of the entities to which it is communicating. For this purpose we generalize the definition of security feature group V. These are the following changes and extensions.

- **Trustworthiness (V)**: the set of features that enables the user and network nodes to inform themselves whether a security feature is in operation or not and whether the use and provision of services should depend on the security feature.
- **Security management (VI)**: the set of features related to monitoring, key management, and virtualization security.

The usefulness of the security feature groups and whether another grouping of security features/mechanisms is more suitable is work in progress. We are considering to make use of e.g. elements from the ITU X.805 security architecture [x805] in a future revision of our 5G security architecture since it

may provide additional granularity in handling security features. In TS33.102 and TS33.401 the use of the security feature groups is limited and they are not used or referred to outside these documents. One potential use of a classification of security mechanisms/features beyond the one in TS33.401 is to somehow check coverage with respect to requirements and threats to make sure security features/mechanisms are in place, in which case components of X.805 could also be useful. However, the use of the above security feature groups for this purpose is very limited.

5 Architecture Enforcement

5.1 Introduction

By *architecture enforcement* we refer to the problem of ensuring that the logical domains and strata of the architecture are properly upheld in an instantiation of a 5G network. For example, it would be rather meaningless to view a Serving Network Domain and a 3P Domain as "separate" if there are no mechanisms in place to ensure isolation and information flow control between the instantiation of these two domains. Similarly, the strata must protect traffic and signalling with them appropriately. In other words, for the architectural components to be useful beyond a mere "abstract thought experiment", they must also be reflected in an operation network.

Architecture enforcement of security and privacy objectives is as always based on implementation and use of mandatory security controls. These controls must be implemented in all domains, i.e. in all 5G entities, networks and network functions and the platforms on top of which they are implemented. Their implementation in the network should provide strong assurance that the security controls cannot be circumvented. The controls should also ensure that 5G entities, networks and network functions are legitimate, can be authenticated and integrity verified and additionally provide privacy and confidentiality of data and users. In terms of our prosed domain based 5G architecture it should be defined how these domains are securely deployed, integrity verified, protected from outside threats and from threats coming from other domains.

The 5G security architecture will reuse components of the existing 4G architecture when appropriate. Examples of such solutions are the 4G security features developed to cope with threats to radio base stations in physically exposed locations (e.g. when the AN Domain is EUTRA) and tampering threats to user credentials in devices (the USIM Domain is protected by the UICC). Other areas will however exhibit new aspects of security e.g. less emphasis on protection at (physical and logical) domain borders (defence-in-depth), the new need for "hardware root of trust" in virtualized settings to ensure legitimate use of resources as well as authenticated point of deployment. The trust model in 5G also becomes more complicated than for 4G as the number of stakeholders in the system and the use of the system for diverse and often mission critical services increases. There will also be new functionality required for network slicing, mobile edge computing, massive IoT, etc.

The coverage of the standardized security controls for architecture enforcement should as in earlier generations of mobile systems cover functionality required for interoperability, secure operation and protection of stakeholders' assets.

5.2 5G Domain Security Architecture Enforcement

The following discussion on how enforce the 5G security architecture we use the definition of domain as given for 5G in section 4. The strata will play a lesser role in this first version of the enforcing architecture. We have chosen this approach since the domains are more directly coupled to the trust model.

Domains should be isolated and only have well defined entry points where controls can be implemented so that only legitimate traffic and signalling can take place. However, as domains in many cases are virtualized the isolation properties and controlled entry points must be enforced by logical means. Clearly, with this usage of virtualization and much more interaction between domains, this is a more challenging task than guaranteeing security in previous generations.

In the following we review the domains in the security architecture defined in Section 4 and comment on their security controls and where new security enablers are needed. The enablers are briefly described in Section 8, for more details we refer to [d3.2].

5.2.1 Infrastructure Domains

As mentioned in Section 4, the Infrastructure Domains contain the “hardware” and this hardware shall in many cases provide a platform for network function virtualisation, network slicing, and mobile edge computing. This will e.g. require new security services trust anchoring of services and verification of platform integrity as discussed in Section 5.2.1.3.

5.2.1.1 UICC Domain

The UICC Domain will in essence be unchanged in 5G as it already in its current form provides all essential security controls like tamper resistance, protected storage for long-term subscriber credentials, use of trusted execution environments, and optionally secure communication over its interfaces. The only new feature in 5G is if the UICC Domain is required to support slicing and thus it needs to provide a shared trusted execution environment. This shared environment will most likely be based on virtualization using a secure hypervisor which can guarantee isolation between slices and provide proof of its integrity. Detailed requirements on such a virtualized environment are not discussed further in this first draft.

5.2.1.2 ME HW Domain

The ME HW Domain may be required to host one or more IM Domains in addition to the ME Domain. Both these domains may be slice aware and thus require to be run in a shared trusted execution environment. The security requirements for at least part of the ME HW Domain will thus be similar to the requirements for the UICC domain.

5.2.1.3 Infrastructure Provider Domain

Most of the security requirements for support of network function virtualisation, network slicing, and mobile edge computing hits the Infrastructure Provider Domain (IPD). The domain must fulfil a number of new security features as discussed in Section 2.1.1. The most prominent ones being providing proof of platform integrity, isolation between slices and control of services deployment, execution and migration. Specific security and trust guidance for NFV can be found in [etsi_nfv]. We note that the following 5G enablers defined in D3.2 [d3.2] will need to be supported: 1) the VNF Certification enabler, 2) the five Security Monitoring Enablers in that certain events and data has to be collected and made available, 3) Network management and virtualisation isolation enablers, 4) support for the Trust Enablers.

5.2.2 Tenant Domain

A first general security requirement for Tenant Domains, which will become very important in 5G, is that Tenant Domains should be bound to the Infrastructure Domain on which they are deployed, where in some cases, more than one Infrastructure Provider Domain may be involved. The form of binding will vary depending on the characteristics of the tenant domain as well as the infrastructure domain employed, typically mutual authentication between a deployed VNF and the underlying HW would be a relevant requirement.

A second general security requirement for Tenant Domains is that they should implement Trust Enabler features enabling users, management systems and network nodes to inform themselves whether a security feature is in operation or not and whether the use and provision of services will/should depend on the security feature. In virtualized environments this will require support for the VNF Certification enabler.

A third general security requirements for Tenant Domains that are slice aware is that they have to be able to provide strong isolation of and between slices.

A fourth set of general security requirements, which already should be in place by current best-practices, is to apply common best practices in IT security, e.g. that domains should implement authentication, authorization and access controls, have encrypted communication interfaces, use separation of traffic and perform systems monitoring and logging.

5.2.2.1 ME Domain

The ME Domain may be slice aware and usually handles an application environment for services hosted by the ME, 5G communication services and the radio interface. Many of the provided services are critical from a security point of view and should be performed in a trusted and isolated environment. The control of the radio interface is of course of particular importance as manipulation of the radio stack may incur serious disturbances and malfunctioning in Access Domain. The ME Domain will support the 5G enablers: Basic AAA and Device Identifier Privacy.

5.2.2.2 USIM Domain

The USIM domain provides the USIM application, other security services or services requiring security, which are hosted on the UICC. The USIM domain is slice aware and thus has to provide isolation of and between slices. As mentioned, this may be supported by virtualization in the UICC Domain. The USIM domain will host support for the 5G enablers Vertical GBA and Group-based authentication. It will also support the enabler Privacy Enhanced Identity Protection.

5.2.2.3 IM Domain

The IM domain supports alternative means to USIM based authentication and will thus benefit from execution in a trusted execution environment and secure storage of security credentials, provided by the ME HW Domain. The IM Domain will implement support for the 5G AAA Security Enablers Authentication of USIM-less devices and Authentication for BYOI.

5.2.2.4 Access Network Domain

The Access Network Domain plays a key role in providing efficient and secure services in native 5G systems. It will support a large number of old and new security enablers and its orchestration plays a key role in the creation of e.g. slices and microsegments even if it in itself isn't slice aware. Trust has to be established with

subdomains in the UE Domain and mediated to the SN domain. Furthermore, the AN domain should support the Security Monitoring Enablers and the Trust Enablers.

5.2.2.5 Serving Network Domain

The Serving Network Domain has traditionally been considered trusted and mainly, from a standardization point of view been protected by well-defined interfaces and protected communication towards other domains. In 5G the SN Domain may be slice aware and rely on VNF and must thus also support Security Monitoring Enablers and the Trust Enablers. Furthermore, the SN Domain will support the Privacy Enhanced Identity Protoection Enabler. In some instances the SN Domain may also have to support the Micro-Segmentation Enabler.

5.2.2.6 Home Network Domain

The Home Network Domain is from a security point of view one of the most sensitive domains as it hosts HSS (AuC and HLR) functionality. Still, it has mainly been up to the operator to implement appropriate security measures to protect its operation. From a 5G perspective it will require at least the same security controls as the SN Domain and support the same Security Enablers. In particular it can be worth mentioning that SS7/IMAP vulnerabilities occurring when used for control communication between HSS and VLR must be handled to counter threats aiming at theft of user credentials and identities.

5.2.2.7 Transit Network Domain

The Transit Network Domain may not be slice aware. It may still play a role in Micro-Segmentation as it may have to provide services with specified performance and quality to support the features wanted by the introduction of Micro-Segmentations . It should in all other respect fulfil the same requirements as the Serving network.

5.2.2.8 3P Domain

The 3rd Party (3P) Domain is a domain which is “untrusted” from a 5G system point of view. Such a domain should only be allowed to interact with native 5G domains via special policy enforcement points having gateway functionality, e.g. application level gateways, and only be able to access well defined services. It is up to the 5G operator to determine the exact security requirements that a 3PP domain should fulfil to be allowed to interact with the 5G system. If e.g. the 3PP domain provides AAA based authentication services, the interactions could be limited to accepting authentication requests using the Diameter protocol.

5.2.2.9 IP Service Domain

As stated in Section 4, the IP Service Domain represents operator-external IP networks such the public Internet and/or various corporate networks and are as such partially or fully non-trusted. No security architecture enforcement is thus in general possible in this domain.

5.2.2.10 Management Domain

A Management Domain will in many cases provide security critical services to the operation of the network and management of all or a subset of the network domains. It may do security management, i.e. ensure that the security services in domains are in place and operational, but it may also do management of the security services including configuration and installation of credentials and keys. Furthermore, a Management Domain may be directly involved in the operation of the 5G Security Monitoring Enablers, the Trust Enablers, and the Network Management and Virtualisation Isolation Enablers.

If the management of slices in a domain are to be delegated to the slice owner, a management domain may also need to be slice aware.

5.2.3 (Additional) UE Domain

The additional UE domain shall fulfil the same security requirements as specified above for UE domain comprised by the ME HW, UICC, ME, IM and USIM domains. Additionally, as interaction with other UE:s may expose an UE to threats not existing when accessing a Access Network Domain, additional security controls may be required.

5.2.4 Compound Domains

The security of compound domains is specified by the requirements on the individual domains comprising it. In certain cases the security controls for communication within a Compound Domains could be relaxed if they e.g. are collocated in a physically secured environment.

5.2.4.1 Slice Domain

A slice domain usually involves slices in more than one of the defined single domain types. Typically, a slice domain is partially managed by the slice "owner". As a slice domain is hosted in other domains the slice domain has to trust the hosting domains and that they provide the isolation and integrity support required. The slice owner should have the possibility to verify this which means that hosting domains and slice management services must support the 5G enablers VNF Certification enabler and/or be given access to some slice-specific API of the Trust Enablers. In general, the slice owner should have the possibity to configure the security offered by a slice.

5.2.4.2 Administrative Domain

As an Administrative Domain is a special case of compound domain, defined not by the presence of special functionality, but by ownership and/or administration it does not imply any additional security requirements compared to those of the individual domains comprising it.

5.2.5 Domain Interactions

Domains can be characterized in terms of the trust different stakeholders have in them, how they are managed, if they are slice aware, and who owns them. These characteristics determine how interactions can take place.

To discuss how interactions relate to trust issues we here adopt a simple but plausible trust model which only considers domains that can be trusted, are untrusted or are semi-trusted by another domain. Further details will be provided in the final version of [d2.2].

If a domain is trusted then all its credentials and security services are accepted. When a domain is untrusted no security critical interactions should be allowed and if a domain is semi-trusted then the scope of offered services is restricted, for example, a Transit Network Domain may be trusted to forward data, but not access it. Trust is often associated with administrative ownership of a domain. A high level view of trust in different domains is a as follows

- 5G-operator domains trust each other as they comply with provisions of the 5G security architecture. However, in 5G setting the need to verify the operator domain authenticity is more outspoken.

- Management Domains are trusted by the domains they manage. It is assumed that the owner of a managed domain ensures that management is performed by a trusted entity, even in the case where management is outsourced.
- 3P Domains with a 5G operator SLA are semi-trusted in the sense that only the agreed services are allowed and subject to use of agreed security mechanisms.
- Access Network Domains that use other technologies than standardized by 3GPP are generally referred to as “Non-3GPP access” (includes for instance Wi-Fi and fixed networks) and are in general considered untrusted. Note however that tight 3GPP integration of e.g. WLAN access may render the WLAN to be considered as “Trusted non-3GPP access”.
- A Slice Domain is semi-trusted by the hosting domain(s).
- Any AN, SN, HN, TN, or Slice Domain trusts any underlying Infrastructure Provider Domain. This is generally necessary simply because these domains ultimately rely on the underlying infrastructure. Of course, this requires prior trust establishment through aforementioned controls such as authentication and authorization and potential use of additional Trust Enablers. Clearly, some additional security controls such as attestation and end-to-end encryption may relax the trust on the semi-trusted level.
- Slice Domains are untrusted by other Slice Domains.

As is seen above most domains in the 5G security architecture are trusted. They may however belong to different management and administrative domains which would require exchange of trust anchors to allow mutual recognition of credentials.

Security measures within a trusted domain are generally left to the domain owner except when interoperability or specific security requirements have to be fulfilled. Interactions between domains have to be protected. In all cases, the involved entities in an interaction between domains should be authenticated and the communication protected:

- If there is mutual trust between two domains, then the domains should be mutually authenticated and exchanged traffic and/or signalling should be adequately protected.
- If a domain is semi-trusted, the interacting domains should be mutually authenticated, exchanged traffic and/or signalling should be adequately protected and it should be filtered in an application level gateway to limit available services. An example of this situation is a 3P domain interacting with a serving network domain via an Industrial Automation Control (IAC) server for industry robot authentication.
- If a domain is untrusted then it should only be allowed to tunnel traffic and signalling through it.

Tenant and infrastructure domain interactions are mainly concerned with trust anchoring in the sense that tenant domains should be able to authenticate the infrastructure domain used and verify its integrity. In the same way the infrastructure domain should authenticate the launched domain and check that it is an authorized entity. Specific security and trust guidance for NFV can be found in [etsi_nfv]. These guidelines are of particular interest for network domain virtualization and slice domains and should be adhered to. As mentioned earlier the VNF Certification enabler is relevant here.

Slice domains should in addition to authenticating the infrastructure domain also authenticate the hosting tenant domain.

5.3 5G Strata Security Architecture Enforcement

As was stated in the introduction in this report we do not treat the strata security in great depth. Here we just note a number of immediate observations.

5.3.1 Application Stratum

For the application stratum there are no or possibly only minor differences compared to 4G. However, it is worth noting that the Application Stratum may depend on slices and e.g. edge computing to deliver special services.

5.3.2 Home Stratum

The Home Stratum will change as it now also must support the 5G AAA Security Enablers and 5G Privacy security Enablers.

5.3.3 Serving Stratum

In the Serving Stratum there will be some major changes as the network domain and transport domain will be virtualized and based on SDN controls for routing. These SDN controls will be key in e.g. implementing slices and micro segments. It also should support the 5G Privacy security enablers

5.3.4 Transport Stratum

In the Transport Stratum there will also be major changes as the Serving and Transit Network domains will be virtualized and based on SDN controls.

5.3.4.1 The Access Stratum

The Access Stratum is a substratum to the Transport Stratum and is in principle the stratum for the control and data mediated in the Access Network Domain and has specific security solutions per radio access technology.

5.3.5 Management Stratum

The definition of the management stratum is new and many new management and monitoring services will belong to this stratum. A number of security controls will be necessary, e.g. for authentication, and authorization of management actions and for support of the Security Monitoring Enablers and Network Management and Virtualisation Isolation Enablers.

6 Trust Model Mapping

In this section, a mapping of the trust model to the architecture will be performed, ensuring that the architecture can properly reflect the trust model. For example, we will verify that the Actors of the trust model can be mapped to a suitable Domain and that the necessary security controls are in place for that Domain when engaging in interactions with other Actors/Domains. However, since the trust model is not complete at the time of writing, this analysis will appear in the final version of the 5G-ENSURE security architecture. We have however performed a preliminary analysis of the Actor-to-Domain mapping. Suggesting adequate support for the trust model in that regard. In this draft architecture, we provide a general discussion on the relation between trust model and architecture.

6.1 Relationship Between Business Models, Trust and Security

Business models and trust are closely related but nevertheless quite distinct concepts. As discussed in Deliverable D2.2 [d2.2], trust is actually one possible response to risk. A stakeholder (the trustor) faced with a threat may choose not to address the threat with security measures (e.g. because it is too difficult or expensive). In such cases they can avoid the associated risk by refusing to engage in activity leading to the threatening event or situation, accept the risk by going ahead and dealing with the consequences of the threat if or when it arises, or transfer the risk by leaving some other stakeholder to manage the risk on their behalf. The first option is a consequence of distrust, but the other two responses involve formation of a trust relationship. If one accepts the risk, one is trusting the system components (which may include other stakeholders) to avoid or prevent the associated threat from arising. If one transfers the risk, one is explicitly trusting another stakeholder to prevent or otherwise manage the risk.

A business model, on the other hand, is a mechanism by which a stakeholder can extract value from an activity. A business model is founded on the idea that by participating, the stakeholder can deliver benefits to others and in exchange obtain benefits whose value to the stakeholder exceeds the cost of participation. A value proposition is a statement of the benefits provided to another stakeholder. In many cases, the benefits received may take the form of monetary compensation by that stakeholder, in which case one can consider the business model as a way of monetizing the value provided through engagement.

Business models and trust are related in three ways:

- the business model depends on stakeholders delivering benefits as expected, and paying (in cash or kind) for those benefits;
- the value provided may in some cases include accepting the transfer of responsibility for managing a risk from another stakeholder;
- the value exchange implied by the business model provides incentives that may affect stakeholder trustworthiness or perceptions of their trust worthiness.

The first of these creates new potential threats, in which actors interact with a system and then fail to provide compensation for benefits received. This happen if a legitimate stakeholder is impersonated by an attacker, leaving the stakeholder to provide compensation for benefits not actually received. It may also happen if a stakeholder simply breaks an agreement and refuses to pay for benefits they did receive.

The best security measure to prevent a stakeholder simply refusing to pay is by establishing a legally binding agreement such as a service level agreement (SLA) defining the rights and responsibilities of stakeholders to each other. By this we mean that the parties can force each other to comply with terms of the agreement through legal action. The terms of the agreements made by each stakeholder then become part of the risk management (i.e. security) measures used by that stakeholder. The ability to enforce those terms may in itself protect the stakeholder from harm by providing financial compensation sufficient to mitigate the harm caused. The fact that terms can be enforced also means they give other stakeholders incentives making their actions more predictable and in line with expectations, thus reducing the likelihood that their actions will be the cause of any threat.

The converse is also true – wherever there is a business agreement relating to the operation and use of a system, there will be an associated trust relationship. In some sense the agreement provides a context in which to manage the consequences of the trust relationship. However, it is possible for trust relationships to exist between stakeholders without any associated business agreement.

6.1.1 Relationship to Architecture

Previous generations of mobile (indeed any) telecommunications networks were designed with no explicit (documented) consideration of trust. Business agreements are needed for stakeholders to operate and use earlier networks, but the architecture itself does not explicitly acknowledge this or explain what the business and trust models are assumed to be. In 5G-ENSURE, we aim to provide the means to understand trust (using the Trust Builder enabler currently being developed in WP3) and thus also to understand what assumptions may need to be guaranteed or at least incentivised through the creation of agreements.

There is in principle no limit to the ways in which value may be generated and compensated. Indeed, there are methodologies for analysing business models from a stakeholder perspective that allow a great deal of ingenuity to be expressed. A good example is the business modelling canvas approach for analysing value propositions, which was introduced by Osterwalder as part of a broader business model generation approach proposed with Pigneur [ost].

Already it is clear (as discussed in Deliverable D2.2) that the range of business models that could be used in 5G networks exceeds the corresponding range in earlier generation networks. This is mainly due to two innovations:

- 5G networks can provide access to a wider range of resources including significant computation as well as IoT data sources;
- 5G networks can use virtualization to provide isolated networks to meet specific business needs in 'vertical' application sectors over a common infrastructure.

The NGMN has defined [ngmn3] several business models and roles for network operators based on these ideas, such as those shown in Figure 7 taken from [ngmn3]:

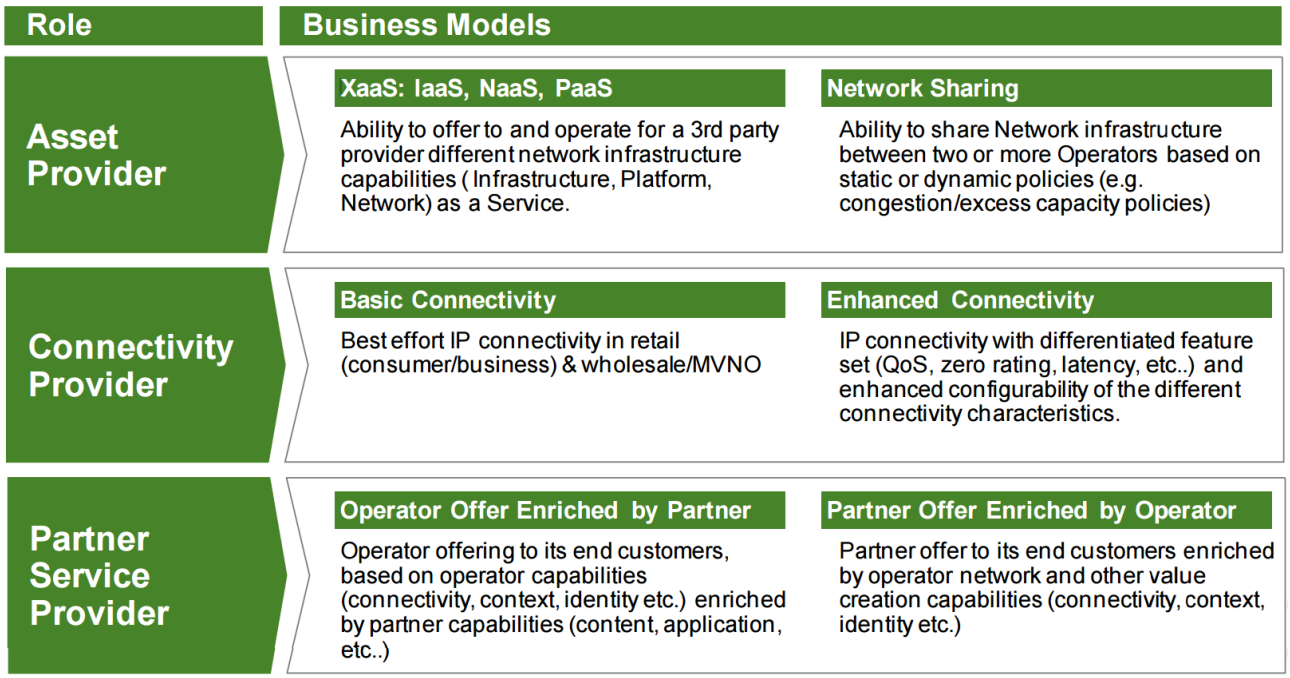

Figure 7: Network operator business models and roles defined by NGMN

Note that the business roles are somewhat interdependent, in the sense that resource provision has no value without connectivity, and partners can only provide added value services on top of connected resources, including their own resources. Note also that some of the suggested business models differ only in the expected value proposition, e.g. the difference between basic and enhanced connectivity is based on the

idea that these will have different value to recipient stakeholders. This idea is not guaranteed to be correct, as the value to a recipient depends on their needs.

The notion of value is not one that can easily be accommodated in an architecture. Value exists in the eye of the beholder, and monetization possible only when the beholder can afford to share some of that value with others involved in creating it. However, architecture can accommodate the notion of trust relationships and business agreements. Given that new business models will probably emerge, our focus is to provide ways to identify trust relationships in a given application of the architecture, so the stakeholders can be made aware of their interdependencies. It is not the role of architecture to specify how they should address these through business agreements, but where they conclude an agreement is needed, the existence of that agreement should be accommodated somehow in the architecture.

Normally business agreements are bilateral (defining the exchange of value between two stakeholders), although the same terms may be used in different bilateral agreements associated with the same system, and sometimes these may be consolidated into a single multipartite consortium agreement. From an architectural perspective, the most flexible approach is to assume all agreements will be bilateral, as multi-partite arrangements can then be modelled as a composition of bilateral agreements [her].

6.1.2 Stakeholder and Trust Models

Work to define a semantic model of assets and threats for use in the Trust Builder enabler is ongoing. At this stage, a partial version is available based on a base set of asset classes shown in Figure 8.

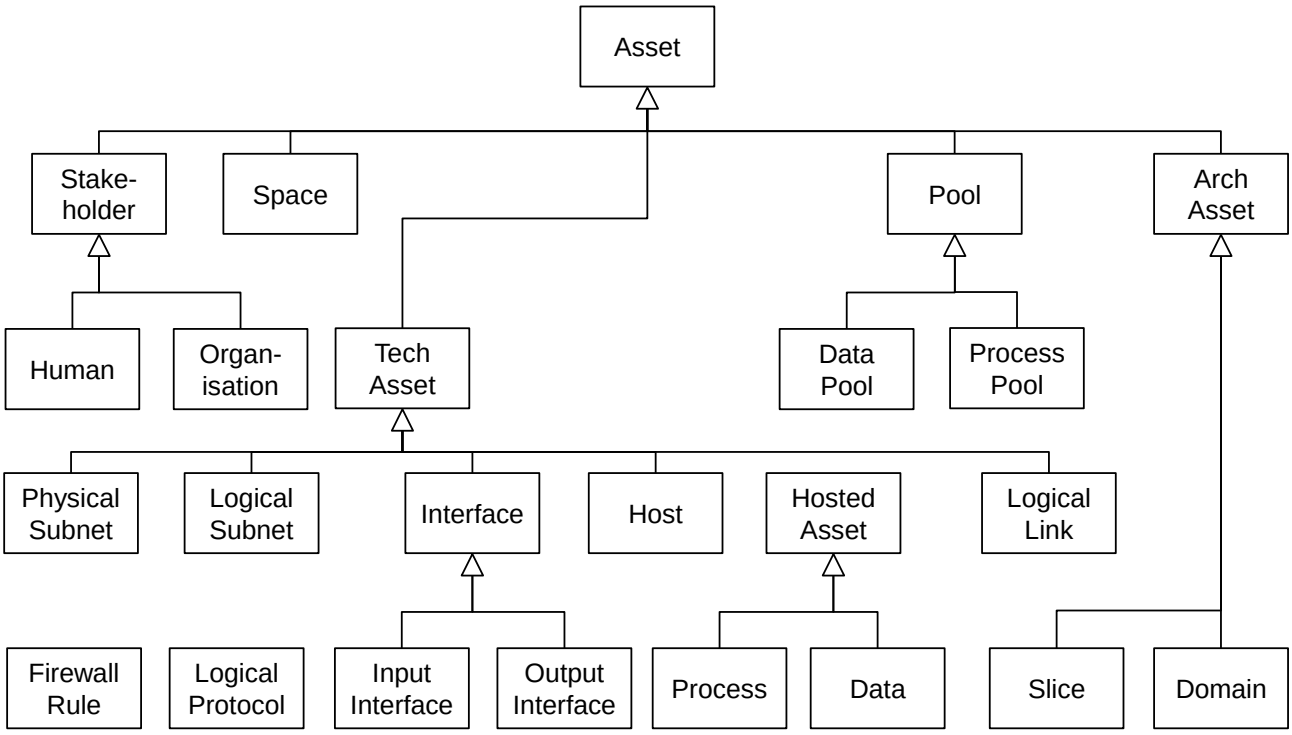

Figure 8: Current base asset classes used in the trust builder

A detailed explanation of all these asset classes will be provided in a future WP3 deliverable. At this stage it is sufficient to focus on a few classes:

- Stakeholder: a class for representing organisational or individual stakeholders in a system.

- Data pool: represents the notion of run-time choices in processing security-relevant data, i.e. data for which there is at least one related stakeholder who is concerned about its security, or on which other security depends.
- Tech Asset: a class representing technology components in a 5G network, including physical and logical subnets, communicating processes, and hosts which are devices capable of communicating, storing and processing data.
- Arch Asset: a class representing architectural constructs with which technology components assets are associated, i.e. network domains and slices.

Stakeholders are considered to be part of the 5G network, thereby treating it as a socio-technical system. This makes it easier to model and analyse stakeholder trust, and the effect on the system if trust is lost or misplaced, as discussed in Deliverable D2.2.

With respect to business models and trust relationships, the interesting question is how stakeholders relate to architectural concepts such as domains and slices, technology components or security-relevant data. Our current understanding of these relationships are summarised in Figure 9.

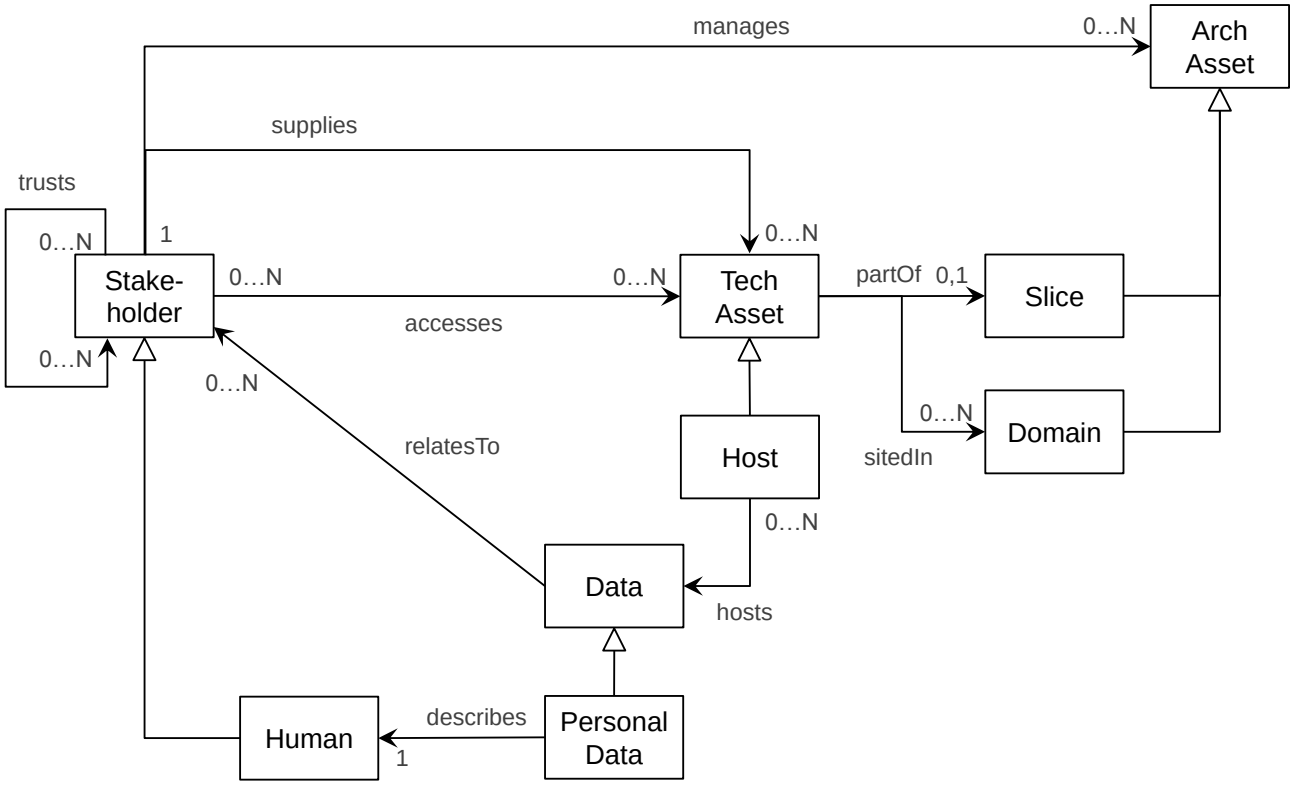

Figure 9: Stakeholder relationships to architectural components

Modelling at this level will allow specific trust dimensions to be analysed in different scenarios, by associating trust relationships with the interaction of stakeholders with and through a 5G network. This should provide insights as to how trust relationships should be managed using 5G-ENSURE security enablers to support the necessary interactions and assurances between actors across domain boundaries. Where appropriate, these may need to be defined in business agreements and use SLA-based monitoring to detect and address breaches.

6.2 Mapping of Actors to Architectural Domains

In this draft, as a first step of mapping the trust model to the architecture, we provide and analyse a mapping of the *actors* of the trust model to the *domains* of the security architecture. In the next revision, using the completed trust model and architecture, we will also map *trust relations* to the *interfaces* between the domains, the *strata*, and the *security feature groups*, required to sustain the *trust relations*. Note that depending on business case, an actor may be mapped to more than one domain. As *primary domain*, we refer to the most typical business case and also list as *secondary domains* which may be relevant in other business cases. The fact that some actors are not mapped to any primary domain may warrant adding domains in the remainder of the work.

Actor	***Primary domain***	***Secondary domain***	***Remark***
Network equip. mfct.		Management domain	Indirectly mapped through the presence of equipment in the Network Domain. The manufacturer (mfct) will in many cases provide management functionality of the products, e.g. updates.
Infrastructure Provider	IP Domain	Management domain	It is likely the infrastructure provider also is in charge of management. This holds most actors below.
Network software (VNF) provider		Management domain	Indirectly mapped through the presence of software in any/all of AN, SN, HN, TN, 3P and Management Domains. Will often provide software management functionality for its software products.
Interconnect network provider	TN Domain	Management domain	
Mobile Netw Op	HN Domain	AN, SN Domain, IP Domain(s), Management domain, Slice Domain(s)	Virtual operators may only have presence in a HN Domain. Commonly, also AN, SN and IP Domains will be provided by the same operator.
Satellite Netw Op	AN, SN and HN Domain	Management domain, Slice Domain(s)	
Network access provider	AN Domain	IP Domain, Management domain, Slice Domain(s)	Since radio networks are probably least likely to be virtualized, the access provider will typically also supply the underlying IP domain
Service provider	IP Service Domain	Management domain, Slice Domain(s)	Over-the-top service providers are Mapped to the IP Service Domain.

User eq. mfct		Management domain	Indirectly mapped through the presence of user equipment in the UE Domain, in particular ME Hardware Domain. Will often however provide management services.
User eq. SW provider		Management domain	Indirectly mapped through the presence of user equipment in the ME Domain.
End-user	UE Domain	3P Domain, Slice Domain, Management Domain	The user may have more or less control over the UE Doman. The user will typically be able to install software in the ME Domain but has (by design) very little control over e.g. USIM/UICC Domain. In some use-cases, the user (e.g. a vertical industry) will also have point of presence in a 3P Domain. Moreover, such users may also have point of presence in a slice and a management domains if the network operator offers users to exercise some control over a slice.
Regulator			No mapping to any domain.

6.2.1 Analysis

The domains of the security architecture seem to well cover the actors involved in the 5G network's "operational" phase, i.e. when the 5G network is delivering connectivity to users via radio/satellite access networks in roaming and non-roaming business scenarios. It seems to also well capture the virtualization aspect of decoupling software from hardware and providing network slicing, as well as the decoupling between different radio accesses and core.

What the architecture captures to less extent is the "out-of-band" relations between, on one hand, operators and on the other hand equipment/software providers and regulators. This will be considered in the remainder if the work,

We can also note that while the architecture well captures the management aspect, there is no "actor" to map to this domain in the case the management is out-sourced. There is likely a need to update the trust model on this point.

These deficiencies will be further analysed and be appropriately addressed in the final version of the architecture.

7 5G Security Design Principles and Recommendations

This section addresses the high level security design principles for the architecture and recommendations as to how to proceed. We take as our basis the set of security building blocks needed to meet the security

objectives, not only in terms of the 5G architecture enforcement (Section 5) but also to address explicit requirements (coming from Task 2.3 “Risk anaysi and requirenets”) and also building blocks related to “common best practices”. We put them into context with reference to the security features. We make recommendations as to how to apply these principles in a 5G architecture. Note that since Task 2.3 is yet in a preliminary phase, this version of the architecture only contains preliminary findings from that work.

7.1 Security Concepts

The high level design principles for the architecture are more general principles which should be applied when designing and deploying mobile systems..

Here we outline the key security concepts that are relevant in the design of most systems. The NIST *Computer Security Handbook* defines the term *computer security* as follows:

> “The protection afforded to an automated information system in order to attain the applicable objectives of preserving the integrity, availability, and confidentiality of information system resources (includes hardware, software, firmware, information, data, and telecommunications).”

These three basic concepts confidentiality, integrity, and availability, commonly referred to as the CIA-triad are key aspects in design of secure systems. Though there are other concepts such as authentication, non-repudiation which are also important characteristics of secure system design.

Furthermore, with an evolving architecture one has bear in mind the backwards compatibility issues, and in particular the design needs to be careful to avoid ‘downgrade attacks’ when attempting to maintain compatibility with older protocols (e.g. key compromise, tunnelled authentication attacks, and cross-layer attacks).

In a mobile network there are a number of layers at which the security mechanism may operate particularly with advent of Internet of Things devices which rely on a plethora or different radio technologies. The deployment across multiple layers requires careful consideration as the system needs to balance a number of factors for an optimum system, such as efficiency, security and compliance with local regulations. Although in certain cases it can improve the security properties and robustness of the system by including similar security services at different layers.

7.1.1 Authentication

Authentication involves the process of providing and assuring identity of communicating entities. Once the authentication phase has completed then the provided identity may be used to grant authorization to utilise certain assigned resources. In the case of the mobile networks the Authentication and Key Agreement (AKA) phase will allow for an authorized user to securely access the mobile network for transport of data for communication purposes. A user’s mobile subscription identity is today defined by their International Mobile Subscriber Identity (IMSI) and an associated 128-bit secret authentication key (K_i), which are usually stored in the USIM on a Universal Integrated Circuit Card (UICC).

- The process of subscriber authentication should employ identity protection preferably through the use of confidentiality and integrity mechanisms.
- The design of the network should be such that unauthorized entities cannot access, nor put a significant stress/load on drain resources, of the core network services.

In the new multi-actor environment of 5G, we of course need to put strong emphasis on always authenticating other entities also more in general.

7.1.2 Confidentiality

Confidentiality provides for concealing of communication content usually through the use of cryptographic algorithms. The use of cryptography is also one of the important techniques to enable privacy for communications and data. Modern mobile networks are now primarily providing for Internet based transport and the main sources of traffic are at the application level which usually employs its own encryption. It could be argued that one layer of encryption is sufficient but it is important maintain a secured transport for a number of reasons. Firstly, one cannot rely upon every single application layer service providing for secure connectivity (e.g. DNS is not yet widely secured), and secondly there are many attacks that may be facilitated by an unsecured transport layer.

There are two basic forms of cryptography. Firstly there is symmetric encryption which uses a single secret key to both encrypt and decrypt communications. The symmetric secret authentication key K_i is used to generate material for use in the AKA protocol which subsequently generate keying material for encrypting the user's communications. . Whilst in today's mobile systems 128-bit symmetric keys are typically used, there are plans to deploy 256-bit keys though no algorithms have yet been specified. Secondly there is asymmetric cryptography which employs two keys one of which may be used to encrypt and the other to decrypt. This approach is also known as public key cryptography, as one of the key pair may be made public, which forms the basis for the use of public certificates which are used in securing key agreement for Transport Layer Security (TLS) and IPsec communications.

The management of keys is a crucial aspect of implementing encryption which for some protocols is separate phase. Furthermore, there are certain properties of key generation and management that can provide for features such as perfect forward secrecy (PFS). The provision of PFS in mobile systems is becoming more important but needs to implemented so that it operates appropriately in conjunction with other services such Lawful Intercept.

- All communications should be encrypted
- It should be possible to provide encryption on end-to-end basis
- The encryption should provide for perfect forward secrecy
- For future-proofness, 256 bit keys should be supported
- Random number sources used in cryptography should be indistinguishable from a truly random source

7.1.3 Integrity

Integrity protection provides for a defence against modification of communication content. This is usually achieved through the use of Message Authentication Codes (MACs), many of which are based upon a range of cryptographic hashing algorithms, though the latest Secure Hashing Algorithm (SHA-3) from NIST employs a cryptographic sponge function. A MAC has the dual purpose of an integrity and authentication check of communication content. Current 3G and 4G standards lack ability to integrity protect the data plane, something which seems necessary e.g. for critical MTC.

7.1.4 Availability

Availability is the property of a system, or a resource, being accessible and usable upon demand by an authorized system entity, according to performance specifications for the system. A number of attack types can result in the loss of or reduction in availability. Some of these attacks are amenable to automated countermeasures, such as authentication and encryption, whereas others require some sort of physical action to prevent or recover from loss of availability of elements of a distributed system. In 5G the

introduction of Software Defined Networks (SDN) and Network Functions Virtualization (NFV) and cloud computing will aid in providing scalable service deployment to maximise availability.

7.1.5 Non-repudiation

Non-repudiation allows either sender or receiver from denying a transmitted message. Thus, when a message is sent, the receiver can prove that the alleged sender in fact sent the message. Similarly, when a message is received, the sender can prove that the alleged receiver in fact received the message. Stronger requirements on non-repudiation arise occur as a need to provide strong liability chains. Non-repudiation could be an important feature when addressing issues relating to inter-operator trust when attempting to tackle potentially fraudulent activities. Networks should provide for non-repudiation functionalities to tackle fraud.

7.1.6 Secure Interworking

Care must be taken when handling back-ward compatibility and interworking with legacy systems. We must ensure protection against "bidding-down" attacks and securely handle e.g. inter-RAT handovers, preventing that legacy systems (with potentially lower security) can lower the security of 5G.

7.2 Standards Based Security

An important principle in designing and building secure systems is to choose security algorithms and security protocols from standardised sources such as those organisations listed below. Clearly new standards will need to be developed but where ever possible they should utilise and build upon existing security standards in preference to developing their own.

7.2.1 ETSI

The European Telecommunications Standards Institute (ETSI) is an independent, not-for-profit, standardization organization in the telecommunications industry in Europe. ETSI produces globally-applicable standards for Information and Communications Technologies (ICT), including fixed, mobile, radio, converged, broadcast and internet technologies. 3GPP is standards body within ETSI.

7.2.2 NIST

The National Institute of Standards and Technology (NIST) is a national laboratory, which is an agency of the United States Department of Commerce, which provides for the development of technology, measurement, and standards. Despite its national scope, the NIST Federal Information Processing Standards (FIPS) and Special Publications (SP) have worldwide impact.

7.2.3 IETF

The Internet Engineering Task Force (IETF) is the main body for standardisation of Internet based protocols and some associated security algorithms. It is an open standards organization, with no formal membership or membership requirements for participation. It is supported by the Internet Society (ISOC) which is an international non-profit organization founded to provide leadership in Internet related standards, education, access, and policy. ISOC has a worldwide membership including both organizations and individuals. It is also the organization home for the Internet Architecture Board (IAB) which provides high level guidance to IETF. These organizations develop Internet standards and related specifications, which are published as *Requests for Comments* (RFCs). IETF has been responsible for standardizing many of the most widely used Internet protocols, e.g. TLS, IPSec.

7.2.4 ITU-T

The International Telecommunication Union (ITU) is an international organization, which is part of the United Nations System, where governments and commercial entities participate to globally coordinate and standardize telecom networks and services. The ITU Telecommunication Standardization Sector (ITU-T) is one of the three sectors of the ITU. The ITU-T mission is to produce standards covering all fields of telecommunications. The international standards produced by the ITU-T are known as *Recommendations*.

7.2.5 ISO

The International Organization for Standardization (ISO) is a global federation of national standards bodies. ISO is a non-governmental organization that promotes the development of standardization and related activities that aims to facilitate the international exchange of goods and services and to develop cooperation in the areas of intellectual, scientific, technological, and economic activity. The work of ISO results in international agreements that are published as International Standards.

7.3 Further Recommendations

Although we have strived to make the security architecture complete, there are obviously issues that fall outside the scope of the architecture itself, e.g. for example how to implement and operate the architecture. In the following, some recommendations are provided.

7.3.1 Implementing Security

Whilst we covered security concepts in the previous section here we examine the ways in which security systems may be implemented. These may just be implemented on general purpose computing entities but increasingly many of these algorithms are abstracted into virtualised services or into specific hardware implementations, or a combination of the two. Making the appropriate design choice as to how to implement security mechanisms can lead to dramatic improvements in security, performance, scalability, and availability.

The Universal Integrated Circuit Card (UICC), or SIM card, is the first example of this trend, which originates from the 2G era, that provides for physical tamper-resistant protection of the user subscriber credentials, the most important of which is the secret authentication key (K_i). Although the external physical interface of the UICC has existed almost unchanged since the 2G days internally the implementation has changed significantly where it has evolved to provide to support for multiple applications, and security improvements like the mutual authentication and use of longer encryption keys. Furthermore, with the advent of the embedded-SIM (eSIM), which allows for remote provisioning of subscriber identities, there are new possibilities particularly for embedded devices which may be too inaccessible, or too small to physically host a UICC card. Although the flexibility afforded by the eSIM has also now attracted consumer manufacturers such as Apple, who have already begun to include it in a number of their devices. eSIM or other forms of trusted execution environments will be key to support large scale IoT and other credential alternatives to (U)SIM.

Trusted Systems in one form or another have now become available in many smartphones. This is largely down to the fact that many now utilise CPUs based upon the ARM chip design which has included ARM's Trustzone for a couple of generations. TrustZone is the marketing name for ARM's Security Extensions, which provides two virtual processors backed by hardware based access control. This lets the application core switch between two states, referred to as worlds, in order to prevent information from leaking from the more trusted world to the less trusted world. Trustzone, and related technologies such as Apple's secure

enclave, may be used for storing critical credentials such as biometric data and credit card details, or for running stronger process sandboxing (e.g Samsung's KNOX). There are range of existing and emerging trusted computing technologies that are making their way into the mobile arena (e.g. Intel's TXT, SGX), which may be employed in the user equipment or the core network. Additional alternatives may be needed for some IoT devices.

Virtualisation has yet to have a serious bearing on end user devices, but it is having major impacts on the core network services infrastructure. There are a range of different applications of Virtualisation in the infrastructure. One major aspect of virtualisation is on the use of virtual machines which removes many of the physical restrictions that may be imposed by running software on specific hardware at a particular location. Virtual machines allow for the seamless migration of software services from one place to other for a variety of reasons including scalability, reliability and redundancy. Another important aspect of virtualisation is the rise of Network Function Virtualization (NFV) which also builds upon Software Defined Network (SDN) to provide for virtualisation of networked resources and services. The combination of these technologies provides for slicing, a key 5G technology. However, whilst these technologies provide a range of advantages they also introduce new security issues which need to be addressed, such as ensuring that the security policy is also implemented in the network management entities.

7.3.2 Design Phases

We here briely discuss considerations for design 5G networks. Note that recommendations here merely serve as examples, no indepth work has been done in the project.

Threat analysis. A comprehensive list of all possible threats against the system needs to be compiled along with the cost of carrying out an attack that can lead to a particular threat. This has begun to be addressed in [d2.3].

Risk analysis. The impact of each threat is measured as discussed in [d2.3]. Estimates are required for both the probability of various attacks and the potential gain for the attacker and/or damage to the attacked side caused by them.

Requirements capture. The results of risk and threat analysis will be taken to formulate the security requirements for the system.

Design phase. The security protection mechanisms are designed in order to meet the requirements. The security architecture is constructed using as pre-existing building blocks such as security protocols or primitives, with new mechanisms defined if necessary. Here the constraints have to be taken into account, and it is possible that not all requirements can be met. This may require re-visiting the earlier phases, especially the risk analysis.

Security analysis. This phase should be performed independently from the other phases so that the system may be correctly evaluated. It may be possible to use automatic verification tools for certain parts of the evaluation but a good deal of the work would need to be performed by experts in the field to properly assess the security of the entire whole.

Monitoring. In order to cover the whole life-cycle of cyber-security threats from design vulnerabilities to alerts triggered by SIEMs, the design of cyber-security monitoring function shall follow a continuous process starting at the design of a network, and looping continuously at run-time in a Deming wheel way. A **P**lan phase consists initially at checking the design of a network. If vulnerabilities are found, a remediation is proposed in a **D**o phase. The remediation is then **C**hecked itself to see if it verifies the QoS contracts. Then,

the remediation is deployed in an **A**ct phase. At run-time, the cycle continues: various sensors and SIEMs trigger alerts, and countermeasures or new remediation are proposed in a **D**o phase, which are checked towards the QoS contracts in a **C**heck phase. Then the countermeasure or the new remediation is deployed in an **A**ct phase.

In terms of component, this chain relies on the following different components:

- Sensors are placed throughout the network in order to send back events or measures.
- Aggregators such as SIEMs correlate the rough information of the sensors and trigger alerts.
- Attack graph engine processes the alert and builds an attack tree where the progression of the threat is clearly shown.
- Remediation and/or countermeasures are proposed in order to stop the attack.
- A verification process checks if the remediation or countermeasure proposed is consistent with the QoS contracts of the Service Level Agreement.
- Then, the remediation or countermeasure is deployed by the security administrator.

Since 5G infrastructure is a metamorphic network, a true issue leads in the changes to be performed on the network in an ad hoc manner at run-time to enforce security functions such as mainly sensors, remediation and countermeasures. Challenges can be either to:

- Place a new security function such as a sensor or a security enforcement point,
- Create, delete or change a flow,
- Move, create, delete or add QoS constraints to a VNF or a VM
- Etc.

Reaction phase. While planning of the system management and operation can be seen as part of the mechanism design phase, reaction to all unexpected security breaches cannot be planned beforehand. In the reaction phase it is vital that the original design of the system is flexible enough and allows enhancements; it is useful to have a certain amount of safety margin in the mechanisms. These margins tend to be useful in cases where new attack methodologies appear faster than expected.

7.3.3 Monitoring

As any IT infrastructure, 5G infrastructures are subject to cyber-attacks which could impact the availability of the services, the confidentiality and the privacy of the users as well as the integrity of the data transmitted. A specificity of 5G infrastructure lies in the stack of virtualized services, each of them depending on (or coming from) potentially from different players. The details of the software used underneath do not have to necessarily be accessible to the upper tenants. All parties of the 5G infrastructure, from the infrastructure operator to the verticals tenant, shall collaborate to endorse cyber-security monitoring in their perimeter of responsibility. Indeed, considering that cyber-security monitoring as one aspect of the service security, and considering that service security performance is part of the Quality of Service [Ie860, section 2.6], the one-stop-responsibility principle developed in [p806] and re-used in [ITU-T E860 section 4], can present the right framework to handle this issue. As a recall, the one-stop-responsibility principle states that "*a single provider is considered as responsible for aspects of the service delivery as seen from a user's point of view ("one-stop-responsibility"). That is, a given user should not need to go beyond the nearest provider for the given aspects of the service. On the other hand, the provider might depend on proper delivery from other providers in order to fulfil its commitments.*"

Vulnerabilities and Threats analysis

Classical IT cyber-attack graph engines are based upon the knowledge of both the software and versions used in order to induce the software vulnerabilities from databases such as the National Vulnerability Database or Common Vulnerabilities and Exposures on one hand and the topology of the network on the other hand. In order to define a node in an attack graph, one needs to use an exploit of a vulnerability, leading to a compromising of a host or a privelige elevation.

5G infrastructures bring new vulnerabilities and ways to describe them:

1. Virtualization in 5G infrastructures brings with it specific threats due to the concentration of the command centres on the SDN controllers, the Orchestrators, the VNF managers and the Hypervisors. If one of those is compromised, it could lead to the compromising of all the entities controlled by each of them, that is to say VNFs and VMs.
2. Organization of the 5G domains in new paradigms such as slices and micro-segments brings new QoS metrics which need to be monitored. Contracts guarantying their respect are at threats when the network is under attack. Monitoring should be oriented in order to fit and serve such new entities.
3. In the context of 5G virtualization stack involving a hierarchy of operators, it is not obvious for the tenant to obtain the software versions used by the infrastructure provider, nor the topology of the network, which are the elements needed for a classical cyber-security monitoring through attack graphs, as mentioned above. Nevertheless, it is probably possible to model malevolent activities at another level such as it is done for Advanced Persistent Threats (APT) where social engineering is involved.

Response

Security issues are detected either at design time according to the knowledge of the topology and of known vulnerabilities, or at run time through alerts sent by products such as Security Information and Event Management (SIEM) which correlates security information at a first level, showing in real-time which software or machine is compromised. Responses to this different kind of detection can scale from a proposal of remediation which will consist of a total re-design of the network, to lighter countermeasures such as cutting a route to an identified attacker. Other intermediary responses are also useful, such as migrating a Virtual Machine, upgrading a version of software or applying a patch.

7.3.4 Orchestration

In order to address business requirements related to the security of the operation, 5G providers will have to operate their network as metamorphic entities which can adapt to counteract on-going threats and changing needs of their customers. The convergence of ICT and telecommunication is on its way but needs new achievements to provide a smooth management framework.

- Virtual Network Functions principle has been standardised to help such a convergence [etsi_mano]. Then, a language like TOSCA [tosca1,tosca2] defines ways to describe objects such as Network Services, Virtual Network Functions, Virtual Links, Connection Points, VNF Forwarding Graphs, Virtual Network Forwarding Paths, Virtual Deployment Units. TOSCA descriptors regroup these concepts to describe the following entities: Network Service Descriptor (VNF, CP, VL, VNFFG, VNFP),
- Virtual Network Function Descriptor (VDUs (CPU, RAM, images), CPs, internal VLs),
- Virtual Link Descriptor (CPs, Type (E-LAN, E-Line, E-Tree…),
- VNF Forwarding Graph Descriptor (VLs, VNFs, NFP (routing policies)),
- Physical Network Function Descriptor (external CPs, linked VLs or PLs).

Such a language enables basically a good description of these entities, as well as VLs and VDUs capabilities and requirements. Nevertheless,

1. Extension of this description shall be made possible to express security specific dimensions to meet Quality of Service requirements,
2. Actual interpreters like Tacker from the OpenStack environment only interpret these descriptors before the deployment of the components, but not at run-time. This is a big limitation for security dynamicity requirements. The schema below presents Tacker's ecosystem.

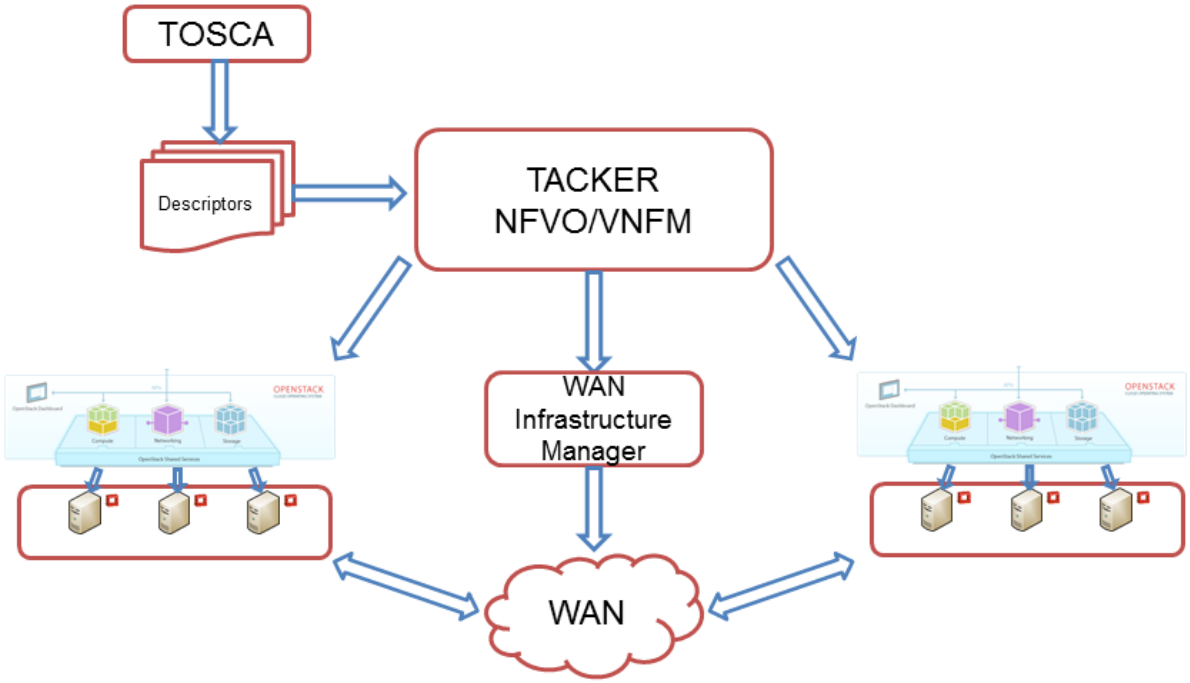

Figure 10: Tacker's ecosystem

In order to counter this limitation for dynamicity, a partial solution can be found in SDN controller's capacity to enable dynamically new routes. Therefore, for example in case of a DDOS attack, a flow can be rerouted dynamically through network functions such as a "clinic centre" in order to benefit from an additional network service. Still, recommendation can be expressed for future work that next generation of VNF Orchestrators shall be able to handle dynamic change request for the deployment of new security functions at run-time.

8 Mapping of 5G-ENSURE Security Enablers

In Section 5, we made a high-level description on how the 5G-ENSURE enablers of WP3 map onto the domains of the architecture. In this section we make a more detailed analysis of the architectural mapping for a few selected enablers. 5G-ENSURE enablers are grouped into 5 categpries:

- AAA enablers, e.g group autyhentication of IoT devices
- Privacy enablers such as improved subscriber identity protetcion
- Trust enablers, presenting users with trustworthiness information
- Security Monitorin enablers, such as decetion of malicious traffic

- Network management & virtualization enablers, e.g. platform attestation

In this draft version of the architecture, we have selected the security monitoring and privacy enablers.

8.1 Security Monitoring Enablers

This paragraph addresses the initial mapping of the Security Monitoring enablers category developed in the context of task T3.4, towards the 5G-Ensure security architecture. These enablers are mainly targeting the supervision of the network in terms of Security Quality of Service reporting and Cyber-attacks detection. All of them basically rely on sensors, and compute the events they receive from them in order to produce reports and alerts.

Starting from D3.2 Open Specifications [d3.2], the analysis of the five Security Monitoring enablers provides a good overview of the infrastructure components they monitor. This section will then map the Security Monitoring enablers category to the 5G-Ensure architecture concepts:

- Domains
- Strata
- Security features

8.1.1 Analysis of the Security Monitoring Enablers

Enabler: Security Monitor for 5G Micro-Segments

The security monitoring for 5G Micro-segments enabler collects information on security related events in micro-segments and infers knowledge by processing and combining these events. The goal of is to enable gaining of real-time awareness of the security situation in micro-segments and detection of some on-going security incidents. The enabler monitors input information from:

- Flow statistic events from Switches,
- Topology events from Network controllers
- IDS events from Traffic inspector
- 5G events from 5G functions such as MME, eNodeB, AAA.

Enabler: Generic Collector Interface aims at collecting logs and events basically generated by the 5G core network components:

- Mobility Management Entity (MME)
- Home Subscriber Server (HSS)
- Packet Data Network GateWay (P-GW)
- Serving GateWay (S-GW)
- Network Function Virtualization (NFV) elements such as NFV Orchestrator, VNF Manager, VIM
- Software Defined Network (SDN) controller
- Authentication, Authorization and Accounting **(**AAA) servers

These events are computed by an Engine. In order to scale to world-wide networks, the treatments of the event collection are performed by engines distributed by layers through the architecture.

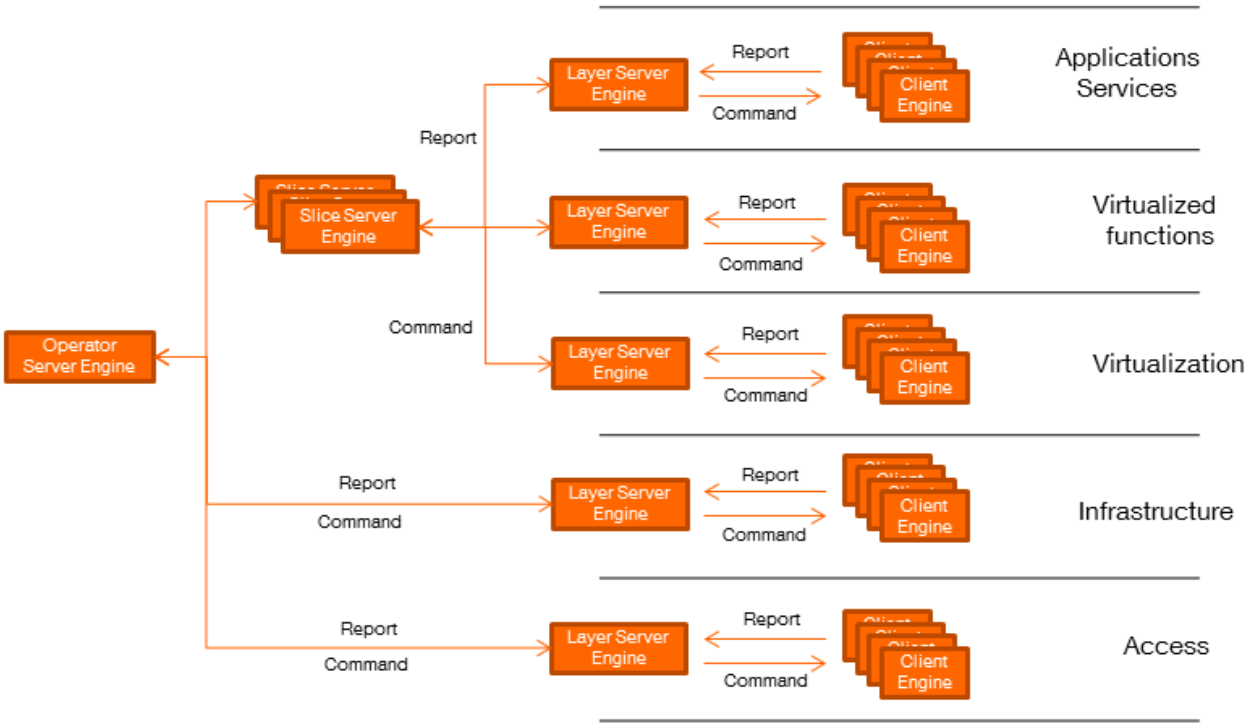

Figure 11: Enabler "Generic Collector Interface"

Enabler: PulSAR (Proactive Security Assessment and Remediation), relies on data inputs from third party components such as a topology scanner to get the topology of the network, a vulnerability scanner to get the version of the software and their known vulnerabilities in databases such as CVE [cve] or NVD [nvd], and a Security Information and Event Management (SIEM) product which sends alerts on on-going attacks. Above that ground need, PulSAR could get information from other Security Monitoring enablers such as the Generic collector or the Security Monitor for 5G Micro-segments, and also from key network and security functions in the architecture:

- Topology information can be given by a SDN controller and all level of VNF management from the orchestrator to the VNF Manager to the VIM to the Hypervisor.
- Hypervisors can provide information regarding versions of the software deployed.

PulSAR targets to produce also remediation and countermeasures propositions which, if applied, induce commands to key network and security functions such as orchestrators, controllers, hypervisors, etc. This draws a loop from these NFVs to the security monitoring enabler, since the monitoring enabler gets information from the NFV, and sends back commands to it.

Enabler: Satellite Network Monitoring focuses on providing pseudo real-time monitoring of logs and alarms on integrated satellite and terrestrial networks and threat detection in these systems. The infrastructure for building the Satellite Access & Transport Networks comprises the following components (see D3.2 Figure 58):

- Satellite Hub: satellite earth station connected to the 5G network.
- eNBs and Satellite-capable eNBs: (traditional eNB improved with a satellite link).
- Different UEs:
  - Satellite Terminals (Ka band): satellite terminal with a Ka band antenna.
  - Satellite Modems: end-user satellite terminal connected to a satellite antenna using a communications satellite as a relay.
- 5G devices.

Enabler: System Security State Repository

This enabler consumes monitoring events from the Generic Collector enabler to provide security information about a runtime system.

8.1.2 Mapping to Domains

The schema below shows the mapping of the Security Monitoring enablers to the 5G-Ensure domains as defined in chapter 4. As explained earlier, the security monitoring enablers are composed mainly by a central engine and get information from sensors or clients. A single color (flashy green) is used to place both the server (a disk) and the sensor (a plain D letter).

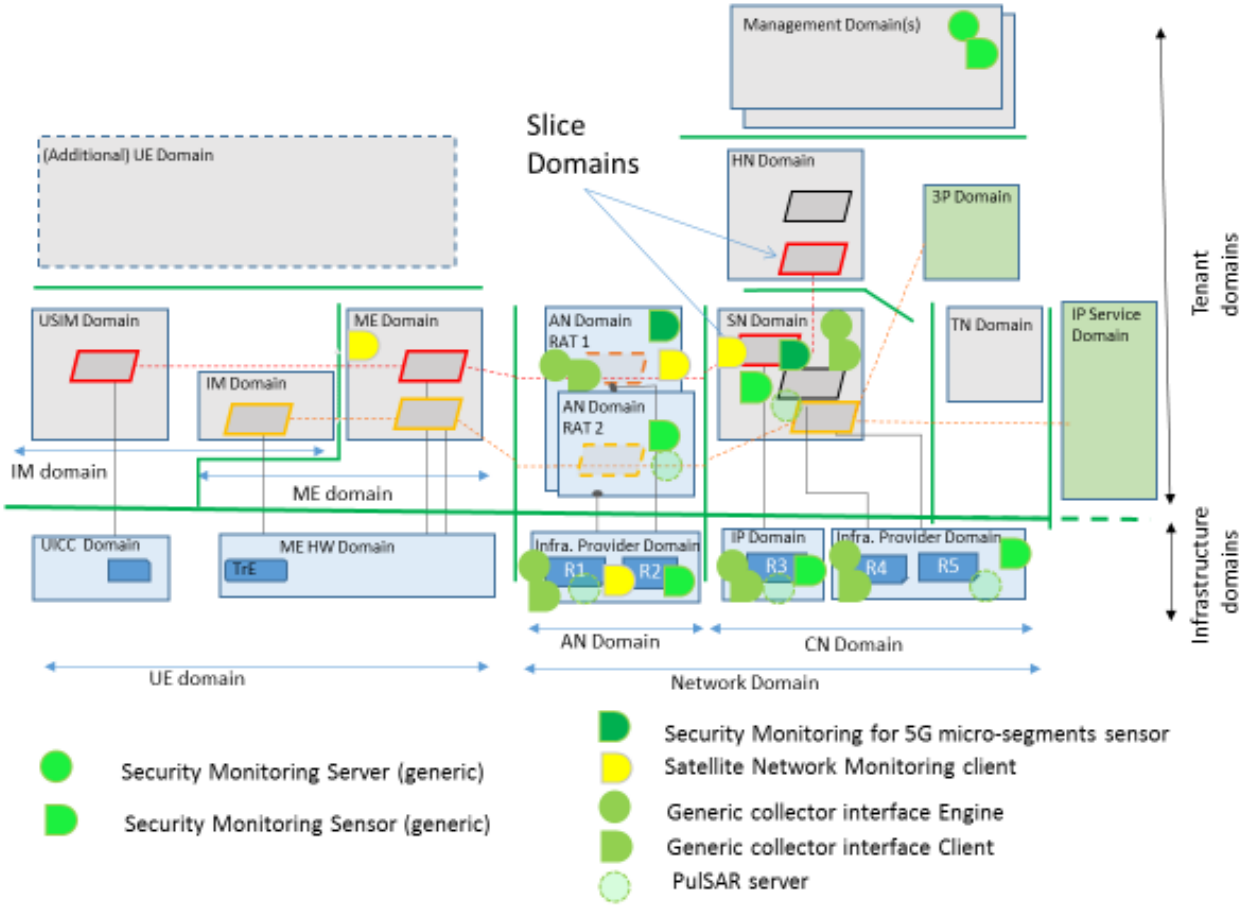

Figure 12: Security Monitoring mapping to 5G-Ensure Domains

Mainly, the security monitoring enabler server is part of the Management domain in order to control the security of the entire network. This is particularly true for PulSAR which even gives orders back to the controlling instances (Orchestrator, SDN controllers, etc.). The sensors or clients of the Security monitoring enablers are mainly getting information across all the different domains of the Network domain, which are the AN Domain and the CN Domain of the infrastructure, and the AN Domain and the SN Domain of the tenant part. Only the Satellite Network Monitoring gets information from the User Equipement.

In this schema, the only domain which is not represented as far as sending data to the Security Monitoring servers is the HN domain from the historical 3G domain, which will probably soon evolve to be fully part of the 5G security monitoring environment.

The generic collector interface server is distributed all through the different domains of the Network domain. PulSAR server, looking after cyber-attacks, is valuable to be installed at all level of the network, from the infrastructure to the tenant domains.

8.1.3 Mapping to Strata

The following schema maps the security monitoring enablers to 5G-Ensure strata as defined in chapter 4. The same iconography is used.

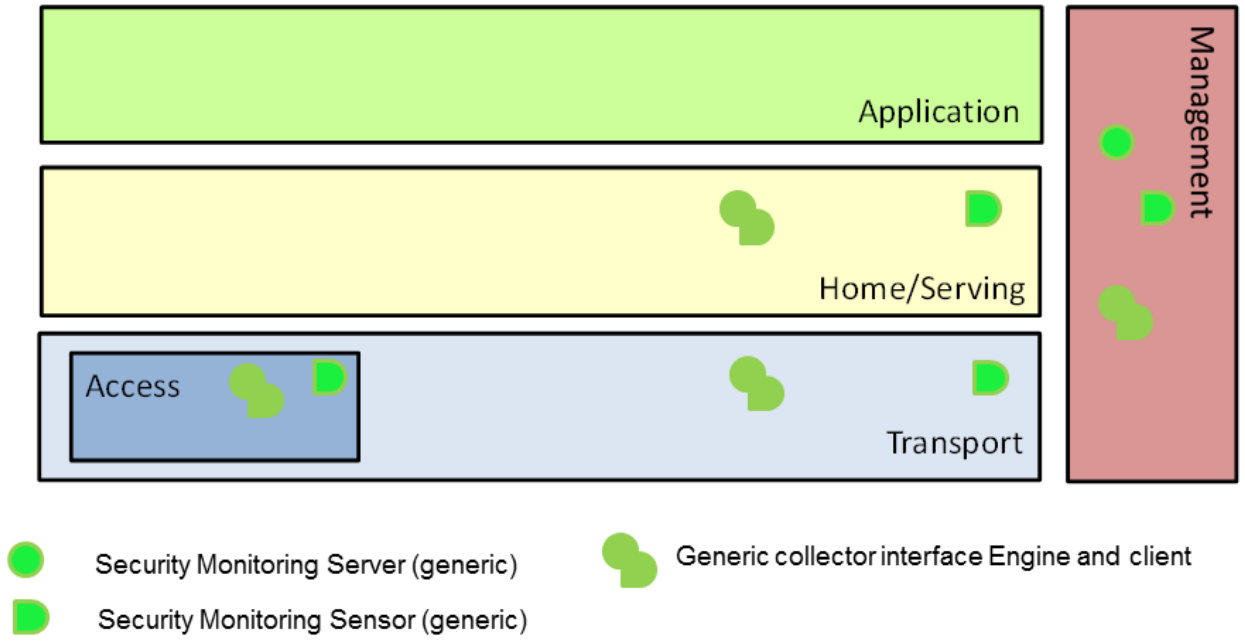

Figure 13: Security Monitoring mapping to 5G-Ensure Strata

The security monitoring enablers target to give an overview of the security quality of service through the network. Therefore, the sensors are deployed through most of the strata, including the Home/Serving stratum, the Transport stratum and even the Management stratum as explained below.

Speaking about the monitoring servers, as mentioned in chapter 5 about the Architecture enforcement, they mostly work as part of the Management stratum. This is either simply because the monitoring mission serves the overall management of the network, or even because they send proposals for remediation and countermeasures requiring changes in the network and infrastructure which should be activated by entities being part of the Management stratum such as orchestrators, etc.

An exception is done for Generic enabler interface enabler which architecture is truly distributed through the different strata. PulSAR, based on the computation of cyber-security attack-trees, could also be deployed locally at an infrastructure provider premises which would be a sub-part of one stratum, in order to prevent cyber-attacks on its network.

8.1.4 Mapping to Security Features Group

The enablers all map to the “Management” feature group as defined in Section 4.2.

8.2 Privacy Enablers

8.2.1 Mapping to Domains

Figure 14 below illustrates the intial mapping of the Privacy enablers to the 5G-Ensure domains as defined in Section 4.

Enabler Encryption of Long Term Identifiers (ELTI) is mainly about IMSI encryption and IMSI pseudonymization, therefore, depending on the architectural choice, this enabler is implemented either in the USIM or ME domains (as far as the UE-side is concerned) and in the Home and Serving Network domains. ELTI has an impact on the Access Network domain, since the IMSI format changes depending on the asymmetric encryption algorithm chosen. It also has impact on the management domain, because the key management and distribution might be a function of this particular domain.

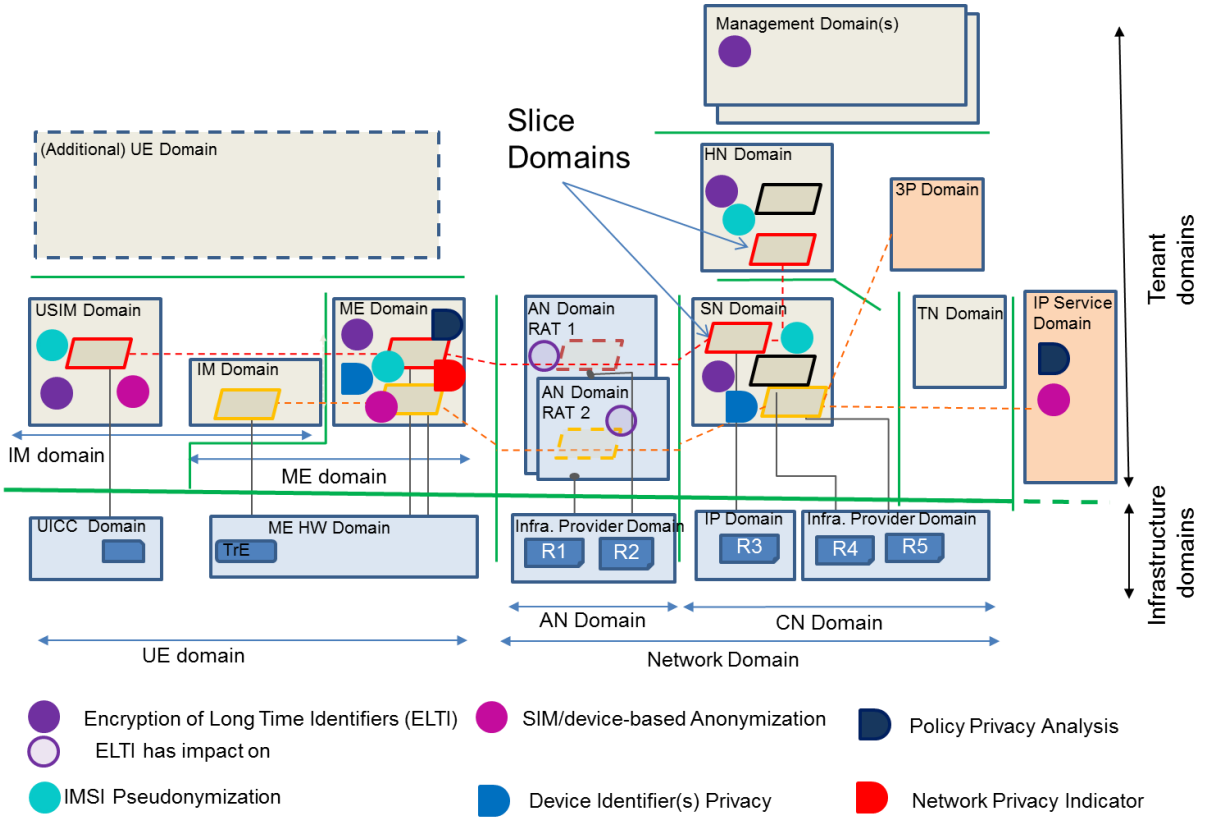

Figure 14: Privacy enablers mapping to 5G domains.

Enabler Device Identifier(s) Privacy lies in the ME domain since it is entirely implemented on the end device. The Network Privacy Indicator is implemented in the ME domain as well.

Enabler SIM or device-based Anonymization and **Polic Privacy Analysis** are more similar at end to end service that the 5G network offers at the application layer, therefore they are included in the ME and IP Service domains.

8.2.2 Mapping to Strata

Figure 15 illustrates the mapping of privacy enablers to 5G-Ensure strata as defined in Section 4. The same colours are used to illustrate the location of these enablers.

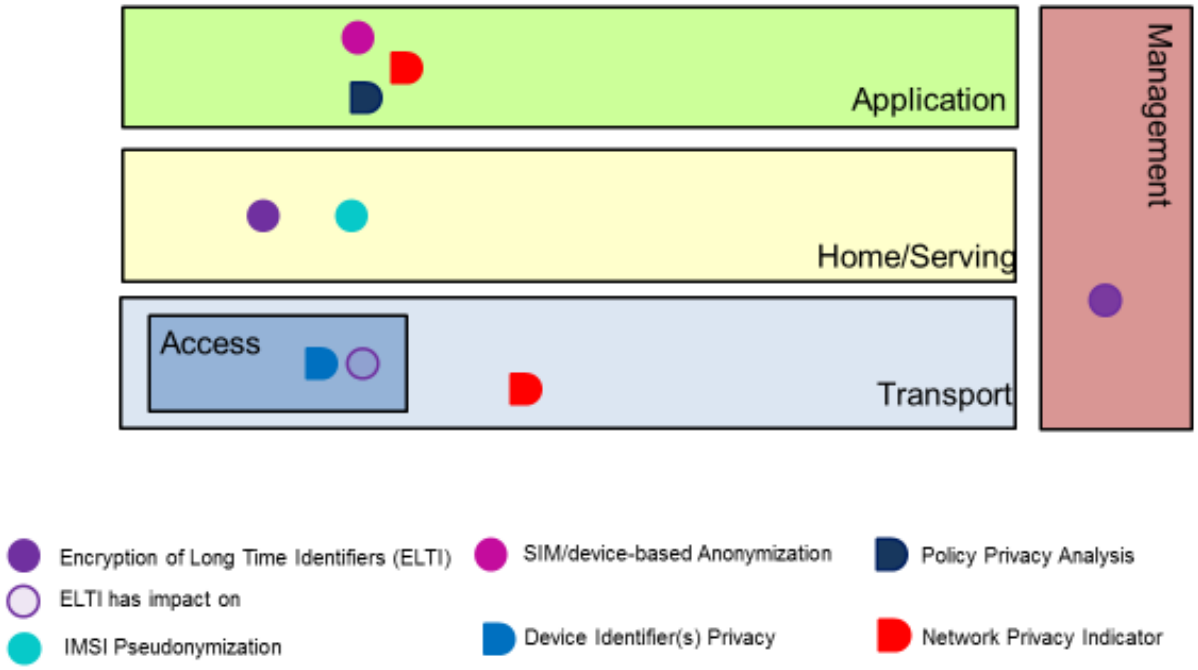

Figure 15 Privacy enablers mapping to 5G strata.

The ELTI enabler mainly targets the Home/Serving stratum, but it has impact on the Management and Access strata as well.

The Device Identifier(s) Privacy can be considered as part of the Access stratum, since it modifies the Detection of Network Attachment protocol.

The Other enablers (SIM or device-based Anonymization, Policy Privacy Analysis and Network Privacy Indicator) are mainly part of the Application stratum, nevertheless, depending on the particular implementation the Network Privacy Indicator may also be part of the Transport stratum.

8.2.3 Mapping to Security Features Group

The ELTI enabler maps primarlity to the Access domain security (I) feature group as it concerns protection IMSI across access, though some implementaions of it also map to the Network security (II) group (protecting IMSI also between MME and HSS). Depending on how keys are stored it may also be associated to User security feature (II) group (with public key in UICC).

The Device Identifier(s) Privacy falls in the Access domain (I).

SIM or Device based Anonymization maps to User security (III).

Finally, Policy Privacy Analysis and Network Privacy Indicator belongs to the Trustwothiness (V) group.

9 Existing Work

As mentioned, there is plethora of security architectures for various purposes. We have here chosen to briefly describe and analyse a few selected architectures which are particularly relevant to the work in 5G-ENSURE.

9.1 3GPP

The security architecture, in particular security features and the security requirements of 3G and 4G networks are defined in [ts33.102] and 4G in [ts33.401] respectively. These two architectures address the 2G weaknesses and provide additional security mechanisms to protect 3G/4G services.

The 3G security architecture defines security features and the security requirements, whereas 4G security architecture additionally defines security procedures in the network. However, the set of reasons behind selecting specific security mechanisms and procedures are not discussed in the 3G security architecture. In a separate study [ts33.821], the rationale and track of security decisions in 4G networks are presented by the 3GPP SA3 group.

Both architectures define a generic security model for 3G and 4G mobile network, however they exclude several factors. For example, while the interfaces between architectural components and their security is defined, end-point (node) security is not considered. Principles together with security objectives are not defined in the architecture documents, however some are discussed in a set of separate documents [ts33.120]. Similarly, 3G and 4G network-specific high-level threats arising from the consideration of security architecture and due to various trade-offs in network performance/availability and regulatory requirements are discussed in a separate technical report [ts33.821].

Definition of trust as an “acceptable level of risk” enables a way of designing secure systems [ros]. In the case of the 3G and 4G security architectures, such a trust level between few network entities is assumed implicitly, and used to derive security protocols and mechanisms. For example, the home environment trusts the serving network by means of service level roaming agreements. In particular, the home environment assumes trust in the serving network to handle subscriber authentication data during AKA protocol in a secure manner.

The 4G architecture extends the trust assumptions made in 3G security as indicated in following Figure 13. The indicated circle demonstrates extended trust between the mobile environment and serving network which is achieved by means of NAS security and key hierarchy mechanisms. However, the 4G security architecture does not define trust levels among network entities explicitly and reasons on why design decisions are at “acceptable level of risk”.

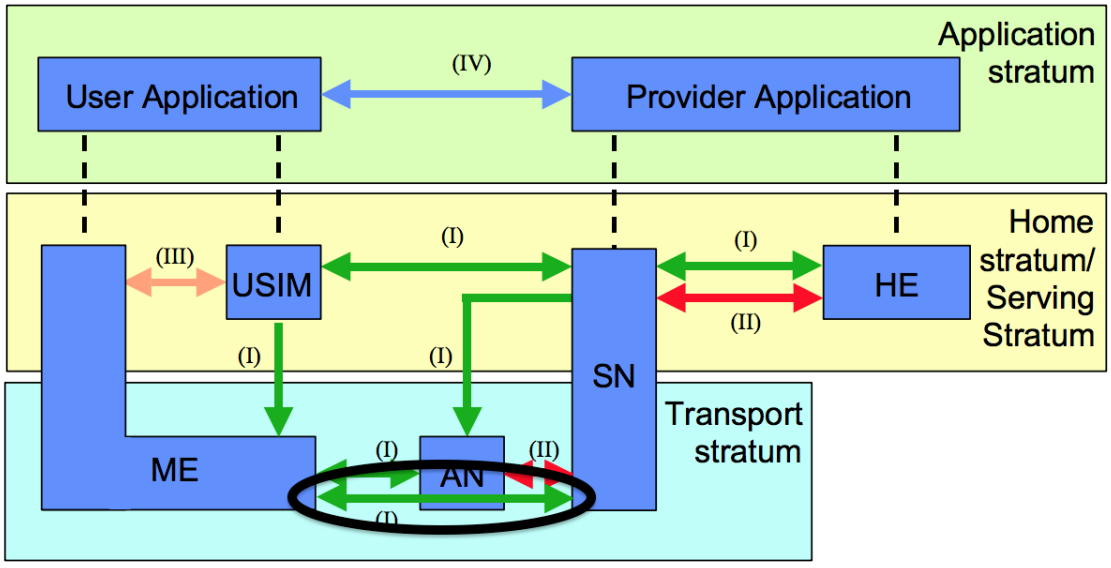

Figure 16: The 4G Security Architecture

Furthermore, implicit trust in user equipment and eNodeB (BTS in 3G) has enabled emerging attacks against both the users and the operator's core network - due to availability of open source tools (2G/3G/4G network software) and low cost hardware. The research work demonstrates practical impact against commercially available devices [nico] and femtocell-enabled mobile network [ravi]. In this 5G security architecture, we revisit the trust assumption carefully and define trust model of 5G network entities to address potential emerging cyber threats to the infrastructure and subscribers.

9.2 ITU X.805

ITU-T Recommendation X.805 "Security architecture for systems providing end-to-end communications" [x805] has been developed by ITU-T SG 17 (ITU-T Lead Study Group on Telecommunication Security) and was published in October 2003. This recommendation defines network security architecture for providing end-to-end network security and the general security-related architectural elements that are necessary for providing end-to end security.

It is based on the concepts of

- Security dimensions (access control, authentication, Non-Repudiation, Data Confidentiality, Communication Security, data integrity, availability, privacy): A security dimension is a set of security measures designed to address a particular aspect of the network security.
- Security Layers (Infrastructure Security Layer, Services Security Layer, Applications Security Layer): they represent a hierarchy of network equipment and facility groupings. Each Security Layer has unique vulnerabilities, and specific threats. For this reason each of these layers must be addressed when creating an end-to-end security solution because at each point the network may be exposed to a new risk, threat or attack.
- Security Planes (End-User Security Plane, Management Security Plane, Control/Signalling Security Plane): A security plane is a certain type of network activity protected by security dimensions. Different security vulnerabilities may exist in each of these planes and each plane along with the three layers must be secured in order to provide an effective security plan.

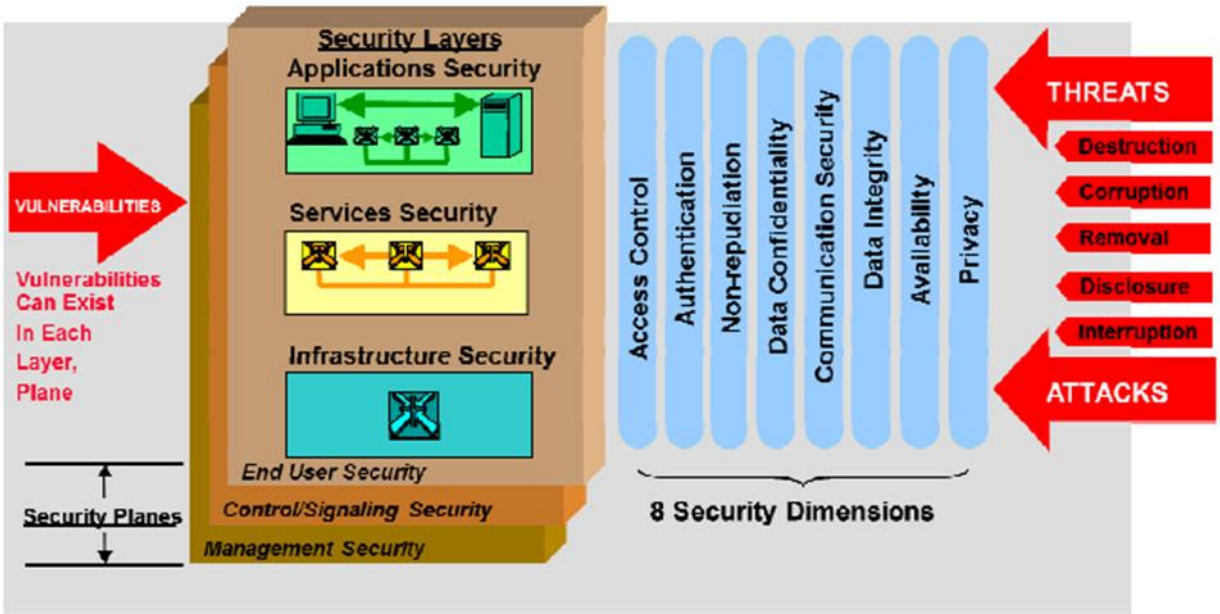

Figure 17: ITU-T X.805 Security Architecture for Systems Providing End-to-End Communications

Then it proposes to apply a security risk methodology to this framework on the basis of ITU-T Recommendation X.800 [x800] which proposes generic threats categories in order to define security objectives and requirements for each cell of the three dimensional table based on the security layer/security planes/security dimensions presented above. For these reasons this approach was also taken into consideration in D2.3 [d2.3].

As a comment on ITU X.805 recommendation, it seems that this framework forces the consideration of all possible threats and attacks to provide comprehensive end-to-end network security. Especially, the differentiation between the security layer and the security planes is very meaningful when it comes to define effective end-to-end security at functional level. Nevertheless, the complexity given by this three dimensional matrix, makes the task quite heavy. Additionally, X.805 does highlight the special importance of (radio) access technology (having specific security issues), nor does it take into account virtualization and thus is further from 5G-ENSURE needs than e.g. the 3GPP architectures.

For this reason, 5G-Ensure security architecture does not directly build on ITU-T X.805 recommendation, although 5G-ENSURE use case description template uses ITU-T X.805 security dimensions category [d2.3].

10 Quality Attributes of the Architecture

This section will contain a quality measurements of the proposed architecture in terms of a set of evidences supporting the claim that the architecture and its security enforcement mechanisms adequately match the security objectives set out in Section 3 as well as more specific requirements coming from Task 2.3 of 5G-ENSURE. However, since neither the architecture, nor Task 2.3 is yet complete, this analysis will be performed only at the conclusion of the work, in the final architecture deliverable. Possible gaps (if any) will also be identified with associated proposals for handling them.

11 Summary and Conclusions

We have presented the first draft of the 5G-ENSURE security architecture. The architecture is based on the already well-established architectures from 3GPP (TS23.101. 33.102 and 33.401). Specifically, it models the network and its security functionality in terms of *domains*, *strata* and *security feature groups*. We have extended the prior 3GPP work to capture all technical characteristics of 5G related to virtualization, multi-access, etc, as well as business model related aspects such as interworking with an external vertical industry's AAA functions.

The architecture design had targeted to be able to provide full alignment with the 5G-ENSURE trust-model (when completed) and, perhaps most important, to provide means for quality assessment in the form of traceability between 5G security objectives and requirements and the final architecture and its features.

However, at this point, it is clear that some important aspects need to be considered in the remaining work. While the domains of the security architecture seem to well cover the operational security and trust aspects of a 5G network, they do not fully capture the relations between the network operator and the equipment/software providers, nor does it capture possible "dynamic" interactions of regulators. We are however confident that these issues can be addressed in the next revision without really changing the fundamentals of the architecture.

In the next revision we will, besides making the needed refinements,

- refine the security feature groups, providing us more precise tools,
- put focus on a more in-depth analysis of the relation to the trust model,
- go deeper into the details of the architecture enforcement and security design principles,
- provide quality evidence of the architecture's soundness and completeness
- further develop mapping of enablers to the architecure.

12 References

[altaf] A. Shaik, R. Borgaonkar, J-P Seifert, N. Asokan, and V. Niemi, "Practical Attacks Against Privacy and Availability in 4G/LTE", In 23nd Annual Network and Distributed System Security Symposium, NDSS 2016, San Diego, California, USA, February 21-24, 2016

[always-on] Tempere University of Technology, Positioning and Location-Awareness in 5G Networks

[android] Android Central, the 'Stagefright' exploit: What you need to know, visited 15 Sept 2016

[bt] BT Group plc, Deutsche Telekom, Ericsson, Hutchison Whampoa Europe, Inmarsat plc, Nokia, Orange, Proximus SA/NV, Royal KPN N.V., SES, Tele2 AB, Telecom Italia S.p.A., Telefonica, Telekom Austria Group, Telenor Group, Telia Company, Vodafone Group, «5G Manifesto for timely deployment of 5G in Europe,» 7th July 2016. Available: http://ec.europa.eu/newsroom/dae/document.cfm?action=display&doc_id=16579

[borgaonkar] Ravishankar Borgaonkar et al., "Weaponizing Femtocells: The Effect of Rogue Devices on Mobile Telecommunications", NDSS 2012

[blackhat] Tom Ritter et al., "I can hear you now: traffic interception and remote mobile phone cloning with a compromised CDMA femtocell", Blackhat USA 2013

[cc] Common Criteria, available at https://www.commoncriteriaportal.org/

[cve] Common Vulnerabilities and Exposures: https://cve.mitre.org/cve/

[cve2] Opendaylight Vulnerable to Local and Remote File Inclusion in the Netconf (TCP) Service

[d2.1] 5G-ENSURE, "Deliverable D2.1 Use Cases," 2016. Available at: http://5gensure.eu/deliverables

[d2.2] 5G-ENSURE, "Deliverable D2.2 Trust model (draft)", 2016. Available at: http://5gensure.eu/deliverables

[d2.3] 5G-ENSURE, "Deliverable D2.3 Risk Assessment, Mitigation and Requirements (draft)", 2016. Available at: http://5gensure.eu/deliverables

[d3.2] 5G-Ensure, "Deliverable D3.2 5G-PPP security enablers open specifications (v1.0)" 2016. Available at: http://5gensure.eu/deliverables

[e860] ITU-T E860 Framework of a service level agreement (06/2002), Available: https://www.itu.int/rec/

[enisa] ENISA, Threat Landscape and Good Practice Guide for Software Defined Networks/5G

[ericsson] Ericsson, "5G radio Access," april 2016. [En ligne]. Available: https://www.ericsson.com/res/docs/whitepapers/wp-5g.pdf

[etsi_mano] ETSI, "ETSI GS NFV-MAN 001 V1.1.1 (2014-12), Network Functions Virtualisation (NFV); Management and Orchestration"

[etsi_nfv] ETSI, "ETSI GS NFV-SEC 003, Network Functions Virtualisation (NFV); NFV Security; Security and Trust Guidance"

[p806] EURESCOM P806-GI project, Deliverable 1

[golde] Golde et.al, "Breaking Band reverse engineering and exploiting the shannon baseband", Recon 2016

[gemalto] The Wired, Gemalto Confirms It Was Hacked But Insists the NSA Didn't Get Its Crypto Keys.

[her] Herenger, H., Heek, R., Kubert, R. and Surridge, M. (2008) "Operating Virtual Organizations Using Bipartite Service Level Agreements", in "Grid Middleware and Services: Challenges and Solutions", Talia, D., Yahyapour, R. and Ziegler, W. (eds), in association with the 8th IEEE International Conference on Grid Computing (Grid 2007), Springer, ISBN: 978-0-387-78445-8.

[huawei1] Huawei PSIRT, UE Measurement Leak Vulnerability in Huawei P8 Phones http://www.huawei.com/en/psirt/security-advisories/hw-459832 , last visited 15 Sept 2016

[huawei] Huawei PSIRT, Several Vulnerabilities in Huawei Honor Routers. Link - http://www.huawei.com/en/psirt/security-advisories/huawei-sa-20160607-01-honorrouter-en last visited 15 Sept 2016

[LTE_book] Dan Forsberg, Gnther Horn, Wolf-Dietrich Moeller, and Valtteri Niemi, "LTE Security", (2nd ed.) Wiley Publishing, 2012.

[ngmn1] NGMN Alliance, "5G security recommendations Package #2: Network Slicing"

[ngmn2] NGMN Alliance, "5G security recommendations Package 1"

[ngmn3] NGMN Alliance, "5G White Paper", Public Deliverable, NGMN 5G Initiative, Feb 2015.

[nico] Nico Golde, "On the Impact of Modiied Cellular Radio Equipment", TU Berlin, PhD Thesis.

[nokia] Nokia Whitepaper, Nokia Threat Intelligence Report - H1 2016

[nvd] National Vulnerability Database: https://nvd.nist.gov

[of] OpenFlow, https://www.opennetworking.org/sdn-resources/openflow/

[ost] A Osterwalder & Y Pigneur (2010) "Business Model Generation: A Handbook for Visionaries, Game Changers, and Challengers" John Wiley and Sons

[p1] P. Langlois, LTE Pwnage: Hacking HLR/HSS and MME Core Network Elements, HITB 2013

[ravi] Ravishankar Borgaonkar, Security Analysis of Femtocell-Enabled Cellular Network Architecture, TU Berlin, PhD thesis

[roger] Roger Rover, "Security Attacks Against the Availability of LTE Mobility Networks: Overview and Research Directions", AT & T whitepaper

[ros] Bill Roscoe et al, "Research directions for trust and security in human-centric computing"

[ruxcon] David Jorm, Redhat, Software Defined Networking Security, Ruxcon 2015

[sdn1] Christian Ropke, "SDN Malware: Problems of Current Protection Systems and Potential Countermeasures"

[sdn2] The PCWorld, "SDN switches aren't hard to compromise, researcher says", http://www.pcworld.com/article/2957175/sdn-switches-arent-hard-to-compromise-researcher-says.html

[tosca1] TOSCA Simple Profile in YAML Version 1.0, Committee Specification 01, 12 June 2016.

[tosca2] TOSCA Simple Profile for Network Functions Virtualization (NFV) Version 1.0, Committee Specification Draft 03, 17 March 2016

[ts23.101] 3GPP, "General Universal Mobile Telecommunications System (UMTS) architecture (Release 13)", (TS 23.101).

[ts33.102] 3GPP, "Technical Specification Group Services and System Aspects; 3G Security; Security architecture" (TS 33.102)

[ts33.401] 3GPP, "Technical Specification Group Services and System Aspects; 3GPP System Architecture Evolution (SAE); Security architecture" (TS 33.401)

[ts33.821] 3GPP, “Rationale and track of security decisions in Long Term Evolution (LTE) RAN / 3GPP System Architecture Evolution (SAE)” (TR 33.821)

[ts33.120] 3GPP, “Security Objectives and Principles” (TS 33.120)

[tr22.891] 3GPP, “Study on New Services and Markets Technology Enablers (TR 22.891)”

[x801] ITU-T Recommendation X.801

[x805] ITU-T Recommendation X.805 “Security architecture for systems providing end-to-end communications”. Available at https://www.itu.int/rec/T-REC-X.805-200310-I/en