White Paper

HardenStance

5G SA Networks Trigger A New Era in 5G Security

By Patrick Donegan, Principal Analyst, HardenStance

Sponsored by

ETSI FORTINET NetNumber SAMSUNG

October 2020

Executive Summary

- Until now, security risk has barely changed at all in the migration from 4G to 5G. Now, with the launch of 5G Stand Alone (SA) networks, 3GPP mitigates some long-standing 4G vulnerabilities to enable much stronger security.
- At the same time, the way the Service Based Architecture 'explodes' the new 5G Core opens up potentially major new vulnerabilities. This requires a fundamentally new approach to securing the 5G Core, including comprehensive API security.
- Operators can communicate 5G SA's new security features to some business users. Communication to consumers is more challenging because the benefit of new security enhancements will only come into effect incrementally over many years.
- Mobile network security cannot depend on 3GPP alone. Operators must apply robust cyber security hygiene and operational best practice throughout their operations.

'Security as Usual' Changes with 5G SA

Until now, 5G hasn't changed the security experience for mobile operators or any of their customers. That's because the first 5G services all leverage the Non Stand Alone (NSA) 5G architecture that connects 5G radios to a 4G core. According to the GSM mobile Suppliers Association (GSA), there are more than 100 5G networks in more than forty countries now. All of them are 5G NSA networks.

Until now, 5G hasn't changed the security experience for mobile operators or any of their customers.

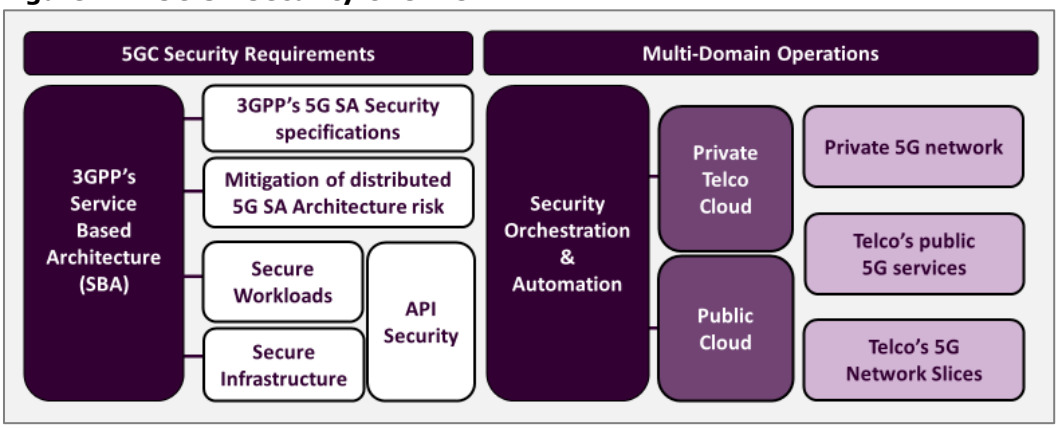
That's about to change. T-Mobile USA launched the world's first 5G Stand Alone (SA) networks in August. China Mobile wants to cover 12 provinces by the end of 2020. Many more operators will launch 5G SA next year. And Vodafone is one of many organisations that has deployed a private 5G SA network – in this case for Lufthansa Technik.

With 5G SA, the security landscape changes fundamentally. This is due to the way 3GPP 'explodes' the traditional mobile core architecture. As depicted throughout this paper, the new 5G Core (5GC) is based on a Service Based Architecture (SBA) or applications services mesh. While this is critical to unlocking a lot of 5G's potential – rapid time to market with new services, network slicing, unprecedented scalability – the new 5G core also introduces some potentially major vulnerabilities. These are common in enterprise IT and public cloud environments but telcos have little experience in addressing them.

Fixing Long-Standing Flaws and Closing Off New Vulnerabilities

This White Paper describes how 3GPP's specifications addresses longstanding flaws in the 4G security model as still used in 5G NSA networks. It highlights key vulnerabilities in the new 5GC, identifying 3GPP fixes where they exist and recommending them where they don't. It also provides guidance on key aspects of customer communications on 5G security and best practise security in day-to-day telco operations.

Figure 1: A 5G SA Security Overview



Source: HardenStance

'Exploding' the 5G Core Introduces New Security Vulnerabilities

Here's how the 5GC specs 'explode' the mobile core and introduce new security vulnerabilities:

As shown in Figure 2, the 5GC comprises many more Network Functions (NFs) than the 4G core. Moreover each NF is made up of multiple microservices for dynamic upscaling and downscaling of services. Rather than communicate via proprietary vendor or telecom sector interfaces, both NFs and unique containerised microservices should only communicate via open APIs. For example, whereas NF to NF interfaces in 4G are telecom interfaces (which most hackers are not at all familiar with), for the 5GC 3GPP has specified open APIs based on the globally popular HTTP/2. The upside is that vulnerabilities will be rapidly identified and closed off by the developer community. The downside is that malicious actors are also very familiar with HTTP/2.

Network Functions must be able to run on any cloud - an operator's private or public cloud, a vendor's own cloud, or a public cloud like AWS, GCP or Azure. This further breaks down the vertically integrated model of vendor-delivered mobile core hardware and software. It signals the end of the integrated vendor model – and with it the end of the mobile core's so-called 'security by obscurity' arising from the dependency on proprietary vendor interfaces that most hackers can't navigate their way around. This also ushers in a new requirement to ensure all NFs are secured against the risk posed by a compromised third party infrastructure – and vice versa. Kubernetes, the de facto telco choice of orchestration platform for the 5GC, has many well-known vulnerabilities that telcos must become masters at understanding and protecting against.

With the 5GC, 3GPP introduces the Network Exposure Function (NEF) to allow direct exposure of 5GC functions to 3rd party applications. This isn't a new concept. The 4G core has a Service Capability Exposure Function (SCEF). However, with the far richer suite of features and services that the 5GC supports, expectations are much higher that the NEF will be extensively used - and with that comes all the associated risk with this kind of exposure.

Six Key 3GPP Security Features of the 5GC

The key security specifications written by 3GPP are captured according to the six high level categories that follow as well as in **Figure 2**. As well as describing the specifications themselves, each category provides the critical business and security context to them.

A totally new feature of 5G security is that the 5G-equivalent of the IMSI is encrypted.

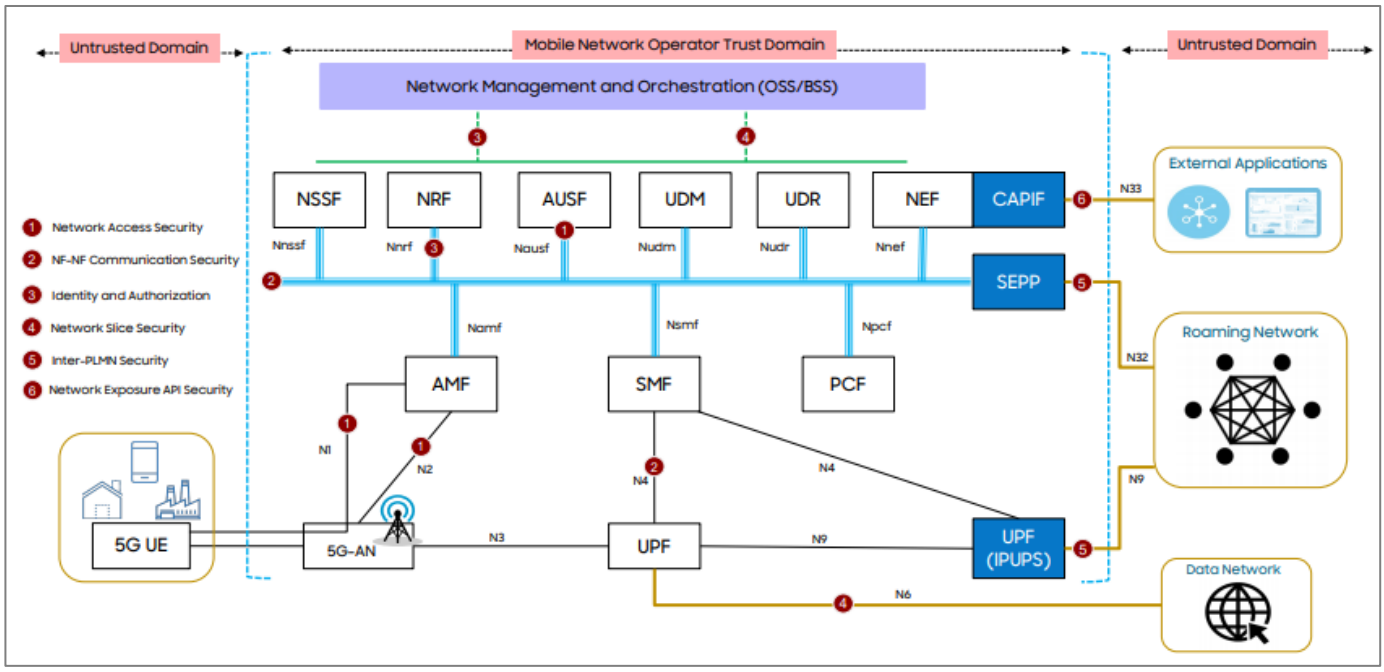
1. Network Access Security

3GPP protects 5G user plane traffic from unauthorized eavesdropping by reusing the 4G air interface encryption which is intended to remain unbreakable for at least the next ten years. Integrity protection is also introduced for the user plane in 5G. This will protect the traffic from unauthorized modifications and hence together with encryption, provides a full-fledged protection for the user plane over the air

The 5GC substantially improves on this by protecting against so-called 'man-in-the middle' attacks using 'IMSI catchers'. In many countries, these devices can be bought legally for a few hundred dollars. They're used as fake 2G, 3G, or 4G base stations to collect data such as whether a specific phone used by a target individual is in the immediate area.

IMSI catchers only work because in legacy mobile networks the International Mobile Subscriber Identity (IMSI) number is transmitted over-the-air unencrypted. A totally new security feature of the 5GC is that the 5G equivalent of the IMSI, the Subscriber Permanent Identifier (SUPI), is encrypted and sent over-the-air as a one-time temporary identifier called a Subscriber Concealed Identifier (SUCI). A second key enhancement in network access security is the new Unified Authentication Framework. This allows the 5GC to serve access requests from Wi-Fi and wireline devices as well as from 5G devices.

Figure 2: Summary of 3GPP's 5GC Security Specifications



Source: Samsung Electronics

2. NF to NF Communications Security

The SBA specifies flat, peer to peer, relationships between Network Functions (NFs) via the HTTP/2-based Service Based Interface (SBI). Due to the fact that most hackers are all-too familiar with HTTP/2 and its vulnerabilities, 3GPP specifies new security requirements that are designed to ensure that NFs only expose themselves to one another securely.

As well as embracing the same requirement as 4G for network domain security using IPSec, 3GPP therefore mandates that 5G SA NFs must also support Transport Layer Security (TLS) encryption to secure the information exchange between them.

3GPP specifies mature, well-established, authentication and authorization standards between NFs.

3. Identity Authentication and Authorization

As well as raising the bar with respect to encrypting the communications between NFs, 3GPP also specifies mature, well-established, authentication and authorization standards between NFs. These are meant to ensure that individual actors and NFs within the 5GC can access only those resources that they are explicitly authorized to have access to.

These apply in two specific contexts:

- The enforcement of policies relating to service requests between NFs. For this, 3GPP has specified the Network Repository Function (NRF) in the role of a 5GC OAuth 2.0 authentication server. NFs in turn have to serve as OAuth 2.0 clients. The NRF issues OAuth 2.0 tokens to NFs only for those specific service requests that a given NF is authorized to make or receive according to the policies set by the operator.
- The telecom operator's operations personnel needing to access NFs as part of Operations and Maintenance (OAM) activities. Here, the operator's Identity and Access Management (IAM) environment must support either OAuth 2.0 or Lightweight Directory Access Protocol (LDAP) to ensure 3GPP-compliant OAM access requests to NFs via the 5GC Element Management Systems (EMS).

4. Network Slice Security

So much of the value in the 5GC is tied up in the capability of the 5GC to instantiate and maintain customized network slices for different use cases, support them across transport, RAN and device, and ensure end-to-end security across those domains.

3GPP provides a lot of the required security layers for network slicing but these will also need to be augmented according to the unique requirements of many different use cases. Beginning with initial slice deployment, 3GPP specifies mutual authentication using OAuth 2.0 between the network slice manager and whatever private or public cloud the slice is being deployed from. In addition, policies also need to be put in place to assure effective isolation of physical and logical networks from one another to ensure threats can't spread between slices. Rate-limiting devices in a given slice should also be considered to protect against DDoS attacks.

Leveraging the Extensible Authentication Protocol (EAP), 3GPP's Network Slice Specific Authentication & Authorization (NSSAA) feature does two key things. At the point of initial connection only devices that support the NSSAA feature, and are authenticated by an AAA server, are allowed to connect to the network slice. Once authenticated onto the slice, and consistent with Zero Trust principles, the NSSAA can then require a device to also request authorization as a condition of accessing any third party application.

5. Inter-PMLN Communication Security

Limitations are in the process of being specified around the extent to which intermediary IPX carriers will be allowed to modify SEPP messages.

Another aspect of the 5GC security standards which represents a major step forward compared with previous mobile generations is the security of control plane and user plane messages between mobile operators - or Public Land Mobile Networks (PLMN) in standards-speak.

The risk of leaving SS7 signaling messages between 2G and 3G networks unprotected has been known about for many years. Attacks on UK and German mobile operators that subjected customers to bank account fraud by manipulating SMS messages via SS7 vulnerabilities have shown that the risk to operators and customers is real now. Hence, thanks to key implementation guidelines developed by the GSM Association (GSMA), leading mobile operators have retrospectively specified and deployed SS7 firewalls as well as Diameter firewalls for 4G.

3GPP has already specified the Security Edge Protection Proxy (SEPP). This protects the HTTP/2 control plane messages between one mobile operator's 5GC and another's, as SS7 and Diameter firewalls protect the interconnection in 3G and 4G. The key thing is that the SEPP is already specified. It's available to operators in advance of them rolling out 5GC roaming agreements at scale rather than them having to wait for these signaling security specs be written retrospectively as in the past.

Ongoing 3GPP Standardisation Work Will Further Harden The SEPP

Ongoing standardisation effort is going into the SEPP to reduce risk still further.

- First, the role of IP Exchange (IPX) carriers in the 5G roaming ecosystem is in the process of being refined. Limitations are being considered around the extent to which intermediary IPX carriers can modify SEPP messages between originating and terminating mobile operator roaming partners. To reduce complexity and risk, the original concept of multiple parties being allowed to modify messages - which both originating and terminating parties would have visibility into as a condition of accepting them - is being refined. The direction in 3GPP now is that only IPX carriers that have a contract with either the sending or terminating operator roaming partner will be able to modify messages – no other intermediaries will be allowed to.
- Second, the Inter PLMN User Plane Security (IPUPS) function has been specified. This enforces GTP-U security on the N9 interface between UPFs of the visited and home operators and can be embedded in the UPF or as a separate NF.

While the new confidentiality protections are supported in the new 5GC, they're not supported in any 5G NSA roaming scenarios. SS7 and Diameter firewalling will therefore continue to be needed to protect 5G NSA sessions against known SS7 and Diameter exploits like location tracking, privacy exposure, SMS interception and fraud.

Increased Home Control

Increased home control protects against unauthorized networks authenticating a customer's device when they're roaming. In 3G and 4G networks, visited networks can be spoofed in order to send false signaling messages to the home network and obtain the IMSI and location of the device for eavesdropping purposes.

5G SA provides increased home control protection by enabling the user device to inform its home network which visited network it is connecting to as well as empowering the home network's Authentication Server Function (AUSF) to report a visited network's authentication result to its Unified Data Management (UDM). If an interaction identifies a mismatch, the home network can overrule the visited network and drop the session.

Some telcos have themselves publicly reported their own high profile data breaches arising from flaws in their API security.

6. Network Exposure API Security

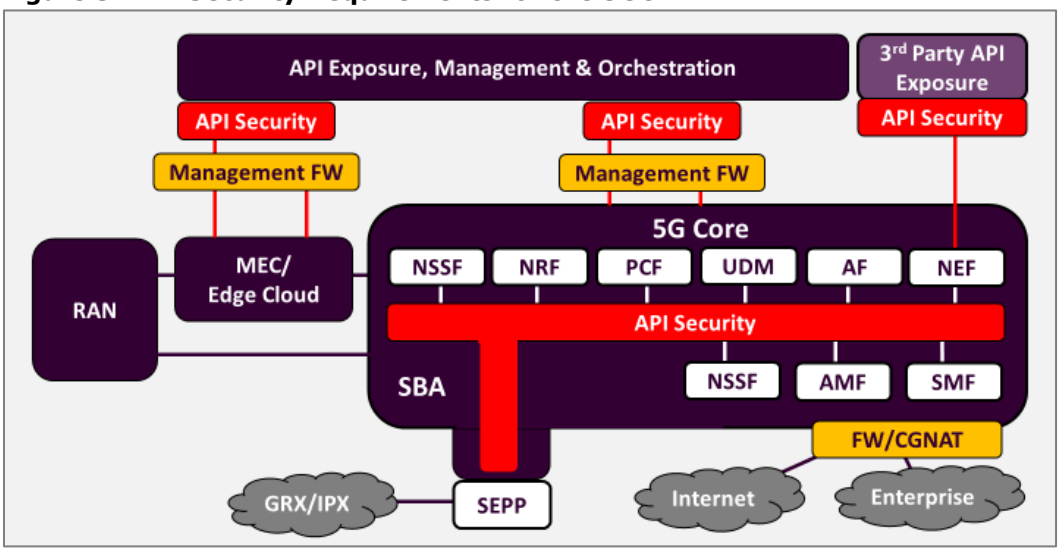
As shown on page 3 and in **Figure 3** the interfaces between 5GC NFs aren't the usual network communication paths – they're APIs. The interface between the 5GC and the management and orchestration layer isn't generic IP communications and things like Simple Network Management Protocol (SNMP) and Secure Shell (SSH) – it's APIs. The interface between the 5GC and the external world of third party applications like IoT platforms, enterprise customer environments and SaaS providers via the NEF is APIs.

Some telcos have themselves publicly reported their own high profile data breaches arising from flaws in the API security of their customer websites. In October 2017, T-Mobile USA notified 2.3 million customers that one such API vulnerability in its website may have compromised their Personally Identifiable Information (PII). In December 2019, Bharti Airtel in India confirmed that 300 million of its customers had suffered a similar fate at the hands of an API flaw in its mobile app.

Say 'Goodbye' to 'Security by Obscurity'

From a security perspective, the design of the new 5GC based on the SBA swaps out a traditional model of vertically-integrated, carrier-grade, telco interfaces with a proven track record of protecting the mobile core through so-called 'security by obscurity'. And it replaces it with an API-driven model with a proven track record of being susceptible to triggering major data breaches in multiple different enterprise environments.

Figure 3: API Security Requirements for the 5GC



Source: HardenStance/Fortinet

At Least as Much to Fear from Benign Errors and Misconfigurations

From a cyber security perspective, the biggest risks introduced by the 5GC are the use of APIs based on HTTP/2 and opening up to third party interactions via the NEF. A mobile core that hackers have traditionally found immensely challenging to penetrate becomes a lot more accessible.

But there's very much more to the change in the risk profile than the threat from cyber attackers. The risk posed by benign errors and flaws - misbehaviour of peering devices, misbehaviour of a configuration – that can take down large parts of the 5GC network is just as great, if not greater. In a 2G, 3G and 4G core environment, these kinds of errors can arise in the context of a relatively low volume of human interactions driven by manual authentications in a closed environment supporting a relatively low number of connections.

The 5GC environment will be fundamentally different. The volume of interactions will be very substantially greater. Interactions will be highly automated between machines whose default setting is to automatically trust one another. The environment will be open to interactions with third parties whose own end points or applications may be misconfigured. And in the MMTC era, the volumes of IoT 'things' connected to the 5GC will be substantially greater. This summer's cyber attack on wearable-maker, Garmin, took the company's services offline for five days. This is only a small taste of potential things to come in this regard.

The risk to the Confidentially, Integrity and Availability of the 5GC from misbehaving devices and applications – whether from malicious threat actors or benign errors – is orders of magnitude greater than any risk to the 2G, 3G or 4G core that mobile operators have faced to date.

Counter-intuitive as this may sound from a security perspective, this choice is entirely necessary. It's the best available way to achieve the goal of a highly dynamic 5GC service creation environment, allow network functions to run in exactly the same way on any infrastructure, as well as make core network functions accessible to third parties. But this unprecedented flexibility does come at the price of increased risk. And that new risk has to be mitigated to harden the 5GC against the risk of major security incidents.

Mobile operators are going to need API security controls that have deep visibility into all API traffic, detect threats like pre-login and post-login attacks, and protect against API-specific DDoS attacks.

To protect the new API-driven environment within the 5GC, 3GPP has specified the Common Application Framework (CAPIF). This provides the mechanisms by which north-bound authentication and authorization, as well as encryption, takes place between NFs and the third party applications domain via the NEF using OAuth2.0 and TLS.

More Effort Needed to Secure an API-Driven 5GC

The foundational security specified in the CAPIF is a good start, but more is going to be needed. Rate-limiting should be considered. For example, this could be used to prevent an AMF from trying to deal with more than a specified number of connections from any one host. Operators can further mitigate risk by selecting a single 5GC vendor for all NFs at launch, albeit that will be at the expense of postponing the greater flexibility and better pricing they will get from a multi-vendor environment.

Moreover, as it's currently specified, the CAPIF doesn't provide comprehensive protection against common types of API vulnerabilities and attacks that can be effective independent of its foundational encryption, authentication and authorization. At key demarcation points - most obviously the NEF and between the 5GC and the management and orchestration layer – the traditional mobile network security model would prescribe a network firewall. But traditional network firewalls aren't API-aware or API security-aware. To secure the 5GC effectively, mobile operators are going to need API security controls that have deep visibility into all API traffic, detect threats like pre-login and post-login attacks, and protect against API-specific DDoS attacks.

Unless different 5GC vendors can agree a common framework for implementing NF to NF security requirements, additional mechanisms are also going to be needed to ensure vendor interoperability within the 5G Core beyond what is currently specified by 3GPP. One approach is to build a TLS lifecycle management microservice into every NF which can interoperate with any Public Key Infrastructure (PKI) or Certificate Authority (CA).

5G SA Security can be Marketed to Businesses…

This White Paper has shown how 5G can provide better security than 2G, 3G and 4G for many legacy voice, text and data services. It has also shown the many measures that secure the new capabilities of the new 5GC as a platform for a new service model, for example in the case of API security and the security of network slicing.

Despite the above, smart operators are not going to be able to go to market with a simple, universal message to their customer base that "5G security is better than 4G". Instead, operators are going to have to craft carefully segmented messages around 5GC security targeting different customer segments.

The wisest choice may be for operators not to communicate anything at all to consumers on 5G security until most or all of their traffic is running on the 5GC.

There is near-term potential for operators to leverage these security enhancements in their 5GC value propositions to business customers. Businesses using a campus network built out using private spectrum and a 5GC, or using a network slice built off a 5GC, can each benefit from all the new security features described in this paper. Hence for these customers, operators can sell 5GC's enhanced security portfolio as hard as they like. They can also differentiate security features that are embedded in the core service from those that can be monetized as a premium, value-added enterprise security service.

…but Consumer Propositions are Trickier

When it comes to consumers, however, the reality of day-to-day mobile network operations won't allow for simple, unambiguous, security messages. For many years to come, most 5G calls will continue to connect over a 4G core, according to the 5G NSA architecture. Hence for many years 5G users will only get a 4G core's security features. It will be years before most 5G calls connect via a 5G SA core, with all added security.

The familiar roll-out scenario in which nationwide coverage takes many years will be exactly the same with 5G SA as with previous mobile generations. But the security implications are different this time. That's because, from a user perspective, there was no noticeable change from 2G security to 3G or from 3G to 4G. With 5G SA, consumers will get very much better security features – but only some of the time, when they're connected to a 5GC. The challenge of whether or not to communicate this nuance to customers – and if so, how – extends into 5G roaming scenarios as well.

Operators therefore need to think very carefully about what to communicate to consumers. Currently, users often see a 5G icon on their phone when they're connected to 5G, signifying a connection to the 5G RAN (and 4G core). Consumers would not respond at all well to two different '5G' icons – one for 5G NSA, another for 5G SA. "5G" and "Secure 5G" doesn't work either as it would risk implying that regular 5G is not secure. The wisest choice may be for operators not to communicate anything at all to consumers on 5G security until most or all of their traffic is running on the 5GC.

Researchers point to Security Flaws in 5G Specifications

Some researchers have made media headlines with research citing potential flaws in the 5G specifications. There is nothing unusual about this. This is part of a normal process whereby researchers identify and publish what they consider to be security flaws. In cases where an issue is acknowledged and deemed to represent a significant enough risk, standards bodies then turn to fixing them.

Operators Can't Live by 3GPP Security Alone

As this White paper has demonstrated, 3GPP specifications and supporting GSMA implementation guidelines are pivotal to how mobile operators need to secure their 5G SA network and services. But mobile operators can't live by 3GPP security alone. The security of the full suite of an operator's 2G, 3G, 4G, and 5G network services is also dependent on how well they carry out routine cyber security hygiene and other aspects of their operations.

In the specific context of the 5GC, there's the security of the underlying and 'exploded' infrastructure platforms that 5G NFs have to run on, whether that's OpenStack, Kubernetes or both. Kubernetes in particular, is new to telcos, and has many substantial security challenges. Most obviously, this includes the simple fact that out of the box Kubernetes allows containers to run as root. That requires a bare minimum of a robust key management solution and Role Based Access Control (RBAC) for risk mitigation.

Mobile operators are going to need security in depth across their operating practises as well as embedded in the services they deliver.

Independent of the specific generation of access technology, operators also need to dedicate a substantial share of their security thinking to shifting security to the left in their development environment – undertaking security checks earlier in the development cycle and more frequently in the operational environment. They also need to master automation of security policies to take advantage of the potential to substantially reduce configuration errors and roll out network-wide security patches at the push of a button.

The 'Soft Cell' Attack exploited an Enterprise IT Vulnerability

Operators must also abide by best practice in enterprise IT security like simple patching and monitoring. When nation state hackers exfiltrated Call Data Records (CDRs) during 2018-2019 in the so-called 'Soft Cell' cyber attack on a number of telcos, they didn't get in via a vulnerability in the 3GPP infrastructure – they got in via a public facing server.

To delight their customers with the 5GC's unprecedented capabilities, mobile operators are going to need security in depth across their operating practises as well as embedded in the services they deliver. A lot of security capabilities are already available and provide what operators need to launch initial 5G SA services. This is a very different operating model, though. Operators need to master the specifications and tools that are already there in terms of security capabilities - and then build on it as they scale up. ■

About the Sponsors

The sponsors of this White Paper are ETSI, Fortinet, NetNumber and Samsung Electronics Network Business.

About ETSI

ETSI provides members with an open and inclusive environment to support the development, ratification and testing of globally applicable standards for ICT systems and services across all sectors of industry and society. We are a not-for-profit body with more than 900 member organizations worldwide, drawn from 65 countries and five continents. Members comprise a diversified pool of large and small private companies, research entities, academia, government and public organizations. ETSI is officially recognized by the EU as a European Standards Organization (ESO). ETSI is a founding partner of 3GPP. For more information please visit us at www.etsi.org

About Fortinet

Fortinet secures the largest enterprise, service provider, and government organizations around the world. Fortinet empowers our customers with complete visibility and control across the expanding attack surface and the power to take on ever-increasing performance requirements today and into the future. Only the Fortinet Security Fabric

platform can address the most critical security challenges and protect data across the entire digital infrastructure, whether in networked, application, multi-cloud or edge environments. Fortinet ranks #1 in the most security appliances shipped world-wide and more than 465,000 customers trust Fortinet to protect their businesses. Both a technology company and a learning organization, the Fortinet Network Security Expert (NSE) Training Institute has one of the largest and broadest cybersecurity training programs in the industry. Learn more at http://www.fortinet.com, the Fortinet Blog, or FortiGuard Labs.

About NetNumber

NetNumber, Inc. brings 20 years of experience delivering platforms that power global telecom and enterprise networks. Our software-based signaling-control solutions accelerate delivery of new services like Private LTE and IoT/M2M solutions across multi-gen networks, dramatically simplifying the core and reducing opex.

These solutions span a range of network types from 2G-3G-4G-5G to future G delivered on the industry's most robust signaling platform called TITAN. NetNumber Data Services are essential for global inter-carrier routing, roaming, voice and messaging. Data powers fraud detection and prevention solutions and enables enterprise B2B and B2C communications platforms. NetNumber multi-protocol signaling firewall, fraud-detection, and robocalling solutions help secure networks against current/emerging threats. For more information visit www.netnumber.com.

About Samsung Electronics Network Business

Samsung Electronics Network Business is a pioneer in the successful delivery of 5G end-to-end solutions ranging from chipset, radio, and core network. The company has been supporting 5G commercial services in leading markets, including Korea and the U.S., where the majority of the worldwide 5G subscribers are currently located, and is also supporting the expansion of 5G in Japan. In addition, the company is rapidly expanding its global footprint to new markets including Canada and New Zealand. Learn more at www.samsung.com/global/business/networks.

About HardenStance

HardenStance provides trusted research, analysis and insight in IT and telecom security. HardenStance is a well-known voice in telecom and enterprise security, a leader in custom cyber security research, and a leading publisher of cyber security reports and White Papers. HardenStance is also a strong advocate of industry collaboration in cyber security. HardenStance openly supports the work of key industry associations, organizations and SDOs including NetSecOPEN, AMTSO, OASIS, The Cyber Threat Alliance, The GSM Association and ETSI. To learn more visit www.hardenstance.com