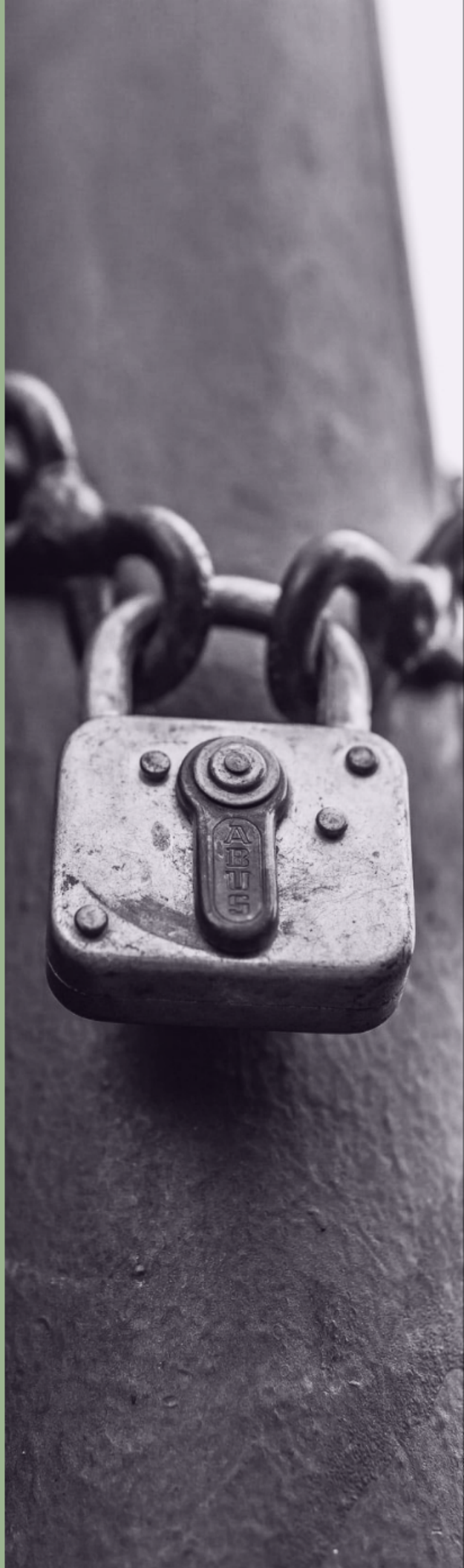
A 5G AMERICAS WHITE PAPER

SECURITY CONSIDERATIONS FOR THE 5G ERA

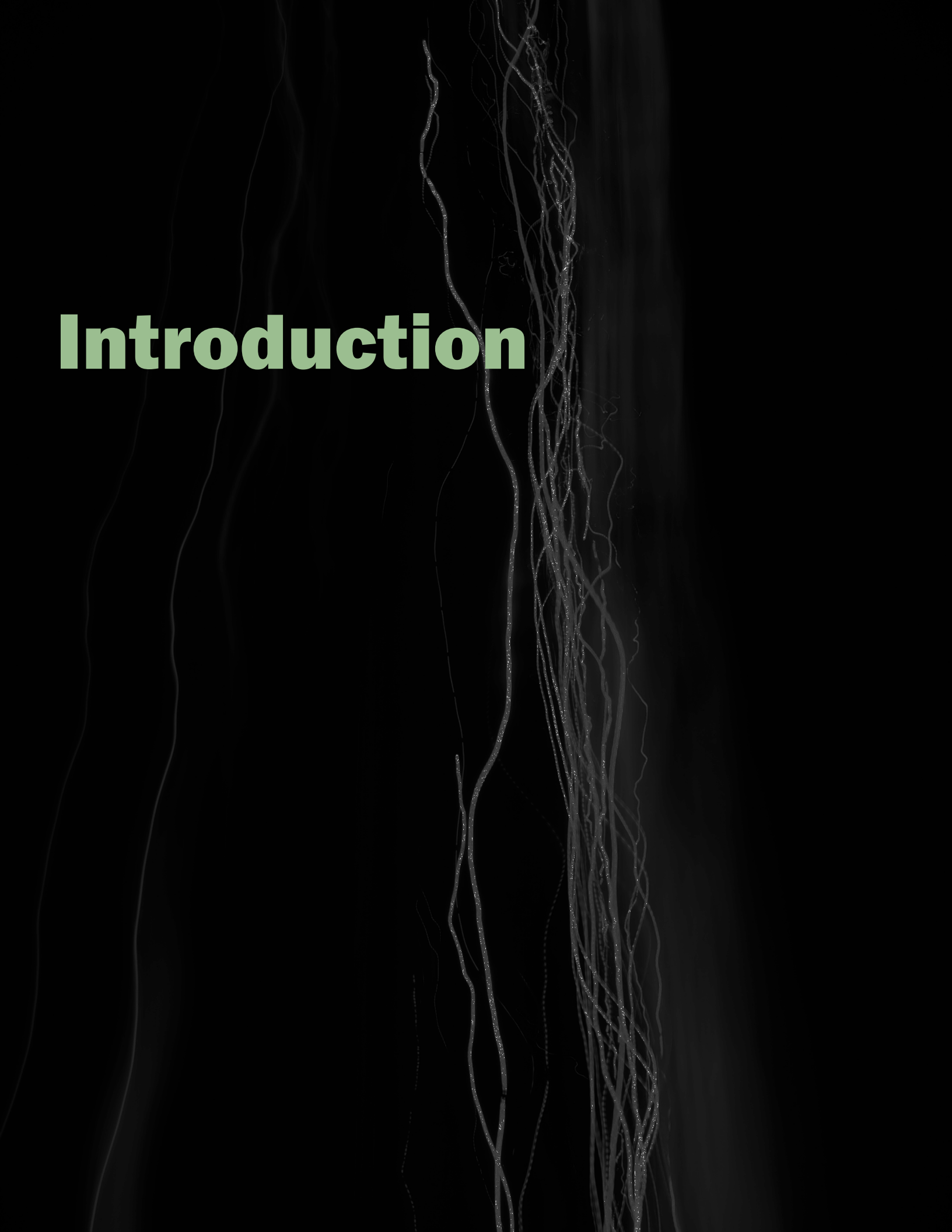
JULY 2020



5G americas

Contents

Introduction



Introduction

5G will likely be one of the most significant technological and societal disruptors of the decade. It promises to deliver higher bandwidth and greater connectivity to enable new services and capabilities that will impact every aspect of our lives. 5G is far more than an increase in radio access bandwidth to the user equipment—it also includes the cloud-based architecture of the 5G core and the capability to support a plethora of things and machines. Use-cases include autonomous cars, integrated smart cities, augmented reality, interconnected social networks and devices, and even 5G-connected cattle for herd management on smart farms. 5G will be ubiquitous.

Proposed 5G architectures are designed to close security gaps from previous iterations of cellular networks, but the pervasive nature of 5G introduces new security challenges outside the traditional space. 5G's attractive, transformative services will likely introduce threat vectors not yet seen or experienced. This paper will look at how 5G differs from other wireless architectures, and what threats, vulnerabilities and attacks are therefore possible. Security considerations will examine various aspects of software, virtualization, automation, and orchestration, as well as Radio Access Network (RAN) considerations. Zero-Trust security, as well as several other techniques, will be discussed to mitigate the threats, and various recommendations will be proposed for security. The paper builds upon and acts as a companion to the previously published 5G Americas whitepapers on security [1] [2].

5G will usher in an age of accelerated innovation, but with that promise comes the inevitable potential for attacks. The telecommunications industry needs to be prepared to defend against these attacks and have mitigation plans in place for current and future attack vectors.

1. Threat Landscapes Throughout Generations

1. Threat Landscapes Throughout Generations

1.1 Why is 5G Different?

5G is the next generation of wireless technology and differs significantly from previous wireless networks. Previous iterations, such as GSM/CMDA (2G) and HSPA/eVDO (3G,) were designed to connect people to people predominantly through voice and text, while LTE/LTE-A (4G) was designed to connect people to the Internet. 5G expands upon this evolution through ubiquitous connectivity of things to people, services, the Internet, and things. To accomplish this, the network is re-architected to utilize software defined networking (SDN) for adaptability, network functions virtualization (NFV) for new services and enhanced capabilities and cloud-native architectures for scalability of resources. The novel network infrastructure enables disaggregation and virtualization, leading to a Control Plane and User Plane Separation (CUPS) with 5G Non-standalone (NSA) and introduces capabilities like network slicing and multi-access edge computing (MEC) with 5G Standalone (SA). 5G SA migrates to a Service Based architecture (SBA) so 5G utilizes a producer-consumer services model instead of fixed functional entities.

1.2 Cloud-Native Architecture: Disaggregation and Virtualization

The 5G architecture takes advantage of cloud-native concepts like self-contained functions within or across data centers (the cloud), communicating in a micro-services environment with all elements working together to deliver services and applications. The cloud-native 5G architecture enables an elastic, automated environment where network, compute and storage services can expand, and contract as needed. Many telecommunications and mobility functions can now be hosted as software services and dynamically instantiated in different network segments. The overall 5G network needs to be pliable and is ultimately designed to be software configurable.

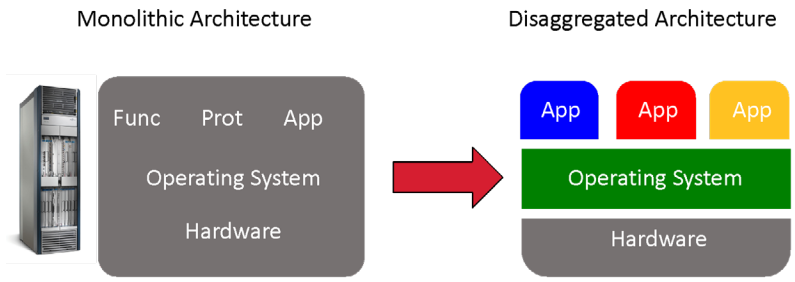


Figure 1: Monolithic versus Disaggregated Architecture

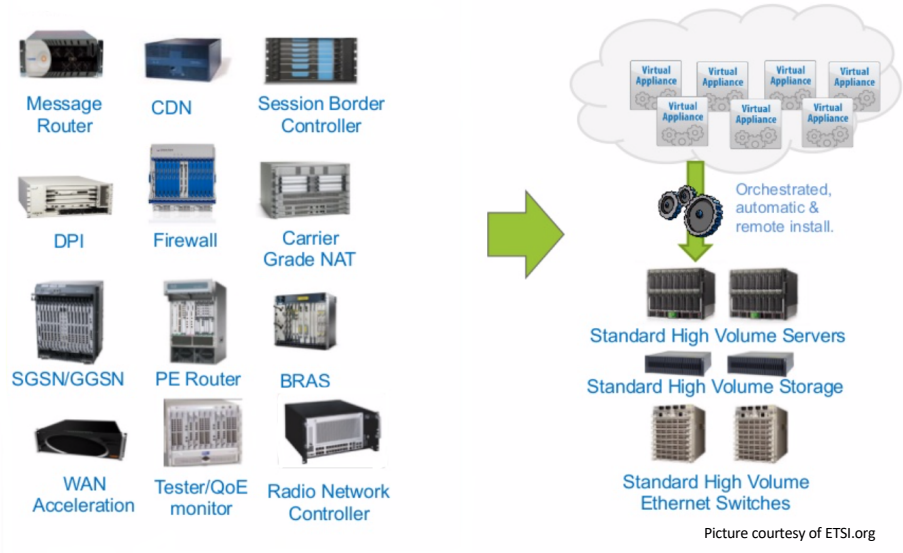


Figure 2: Evolution from network appliances to virtualization

Control plane and user plane separation is the concept of disaggregation that allows these two planes to exist on separate devices or at separate locations within the network. Separating the control plane from the user plane allows the two planes to scale independently, without having to augment the resources of one plane when additional resources are only required in the other. This separation allows the planes to operate at a distance from each other; they are no longer required to be co-located, further improving scalability.

In the Radio Access Network (RAN), migration to the cloud allows the disaggregation of the Remote Radio Unit (RRU) from the Baseband Unit (BBU). By separating these functions, it becomes possible to create a pool of BBU resources that support several distributed RRUs. Doing this allows for more efficient use of RAN resources. However, it also creates challenges, such as the need for fronthaul connectivity between the RRUs and the BBUs, where fronthaul demands high bandwidth and low latency. 5G addresses this fronthaul challenge through a novel architecture that defines splits at different locations in the RAN, where RRUs connect to distributed units (DUs), subsequently connecting to centralized units (CUs). This is referred to as the CU/DU split and introduces the concept of midhaul. How and where an operator decides to place the RRUs, DUs and CUs are driven by trade-offs between bandwidth requirements and the ability to centralize resources.

Another major difference between 5G and previous generations of wireless networks is the radio spectrum being used. 5G requires much larger amounts of bandwidth, and therefore the quantity of spectrum has increased, and its locations are different as well. Some frequency ranges remain below 6 GHz, while a move into the millimeter wave (mmWave) range provides access to much larger amounts of contiguous spectrum. One disadvantage of the higher frequency ranges is the shorter distances their signals can propagate, and their inability to penetrate fixed objects (including rain fade). These limitations impact the quantity and location of base stations needed for proper coverage. This can be solved with a larger number of smaller base stations, i.e. Small cells, Micro cells, Femto/Pico cells. A larger quantity of sites needing interconnection produce a more distributed architecture, described later in this section.

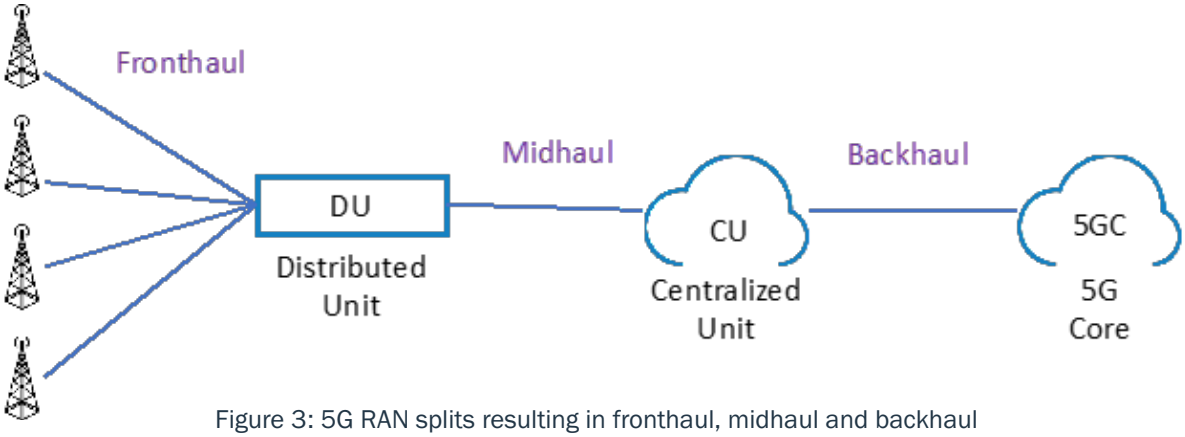


Figure 3: 5G RAN splits resulting in fronthaul, midhaul and backhaul

1.2.1 Network Slicing

Network slicing refers to the technique of isolating the end-to-end performance of a portion of the network compared to another. It involves all three domains of 5G: the RAN, the transport network and the 5G Core. Network slicing is different from Virtual Private Networks (VPNs) in that VPNs are overlays on top of physical network resources whereas slicing partitions the resource. This separation goes all the way down to the user equipment (UE), whether that be a handset, fixed-wireless access point or IoT sensor. Network slicing delivers the stringent characteristics that 5G offers (eMBB, mMTC, URLLC) for specific subsets of users, operators, or applications. Use-cases taking advantage of a network slice vary, but examples include consumer smartphones, industrial automation, autonomous vehicles, law enforcement or emergency/first responders.

1.2.2 Distributed Architecture

5G will have a much more distributed architecture than previous wireless versions. A distributed architecture is needed because: the variety of frequency ranges used to deliver high bandwidth and the density requirements of 5G require many more base stations or small cells; and 5G latency requirements necessitate that some application processing occurs much closer geographically to the user. Other reasons that contribute to the distributed architecture of 5G are:

- Densification of Radios: the need for increased coverage and better experience per device will mandate an increased number of radios.
- Increased Centralized RAN deployment: with a Split CU and DU architecture being defined in 3GPP 5G specifications, there will be an increase in centralization of baseband processing and radio control functions in many locations.
- Introduction of Edge Compute: the distribution of compute power closer to the edge to provide better quality of experience and to fulfill new use-cases.
- Introduction of New Applications: in addition to standard voice and broadband data, 5G introduces new applications, including those based on D2D (device-to-device) communications and IoT, that will drive new traffic patterns for data and control traffic.
- Expected Bandwidth Increases: With the opening up of additional spectrum in the mmWave band, Band 42 (3.5GHz) and C-Band for 5G NR, coupled with antenna techniques such as Massive MIMO, the amount of transport bandwidth is going to increase at the Fronthaul.
- Deployment of Deterministic-Behavior URLLC Service: To handle low latency applications and guarantee their upper bound latency limits with fixed jitter and low packet error loss rate, deterministic networking packet transport technologies will become important.

1.2.3 Service Based Architecture

Each generation of wireless networks contains several services imbedded in the system. These services perform various network functions, from validating if a subscriber can connect to the network to managing the flow of the user equipment traffic. 5G introduces the concept of an SBA, one in which the individual functions or services are virtualized and "cloudified." As a result, instead of a rigid path through specialized equipment—and its physical arrangement based on how it's actually connected—an SBA allows all the functions to communicate with each other via APIs across a packet-based fabric.

1.2.4 Traffic Patterns

The nature of 5G traffic patterns will be unpredictable and ever changing. With a distributed architecture composed of virtualized elements, traffic will flow between services, elements, functions, and devices. Such a distributed architecture may introduce additional attack points, but may also make attacks harder to execute. With a disaggregated packet core, parts in one location will be communicating with parts in another location. The packet core itself will become geographically diverse and more resilient, e.g., implemented by a primary packet core and a backup packet core. 5G enables "any-to-any" communication, for edge-component to edge-component traffic flows.

1.2.4.1 Software

One of 5G's goal is to make the network flexible and programmable. This can be achieved by the virtualization of many functions and services previously implemented in hardware. Software can take advantage of open-source software modules, but may have some associated vulnerabilities.

Vulnerabilities can be reduced when open-source communities thoroughly test software to discover and correct most of its flaws. Potentially greater vulnerabilities exist if a software supply chain does not have safeguards in place. If the origin, reliability and integrity of the open-source modules are unknown, it can introduce flaws and vulnerabilities. Dependency is another conflict; if a software module relies on other pieces of code, then the sudden lack of access to that dependent code could be problematic.

1.3 Threats, Vulnerabilities and Attacks—Pre-Full 5G Standalone (LTE and 5G Non-Standalone Networks)

The first commercial launch of LTE was in 2009 [3], and at that time, cybersecurity or national security was a crucial concern for operators, regulators, and/or government agencies. In the past 11 years, the cybersecurity threat landscape has evolved into a critical focus for national security agencies, as cyber threats can be executed against a nation's critical infrastructure, defense systems, global financial systems, etc.

5G Americas has previously reported on many known LTE threats, vulnerabilities and attacks which will be briefly described in this section. It's important to not lose sight of these known LTE security vulnerabilities; LTE networks and 5G NSA networks based on LTE core networks will continue to operate in the American operators' networks for years to come. Unlike in 2009, the North, Central and South American operators are now threatened by well-funded and patient nation-states. These entities have militarized technology and communication systems for use in a cyber warfare attack against critical infrastructure, military response capabilities, and could cause catastrophic global market disruptions. Known LTE vulnerabilities and weaknesses could also be targeted at key politicians, national security officials, national intelligence agents, local/regional/national law enforcement agents, corporate leaders, celebrities, and even specific defense industry engineers. Sensitive information obtained—passively or actively—could be exploited by another country's counter-intelligence organization for a variety of purposes, including bribing of individuals for monetary information, or additional sensitive information gathering (i.e. spy recruiting).

Early 5G commercial launches are leveraging 3GPP's Release 15 Non-Standalone 5G specifications, meaning these early 5G NSA networks are required to use the LTE control plane protocols and the LTE Evolved Packet Core (EPC) network (Figure 4). Initial 5G NSA launches will deliver only Enhanced Mobile Broadband (eMBB) [4], which is one of the three ITU use cases for 5G (mMTC and URLLC being the other two). For initial eMBB service, any LTE threats and vulnerabilities will also exist in the 5G NSA network. Even after the operators upgrade their cell sites with 5G radios (gNBs) using the NSA architecture, some of these 5G NSA cell sites may operate as 5G NSA sites for years after Standalone 5G architecture is commercially launched. NSA architectures are expected to live alongside each other for a considerable period. The initial 5G SA commercial networks will leverage 3GPP's Release 15 5G Standalone specifications which introduce a new 5G Core (5GC) network and some new protocols that will mitigate some of the known LTE protocol weaknesses. A 5G SA network will leverage neither the same LTE control plane protocols nor the LTE EPC network (see Figure 5.)

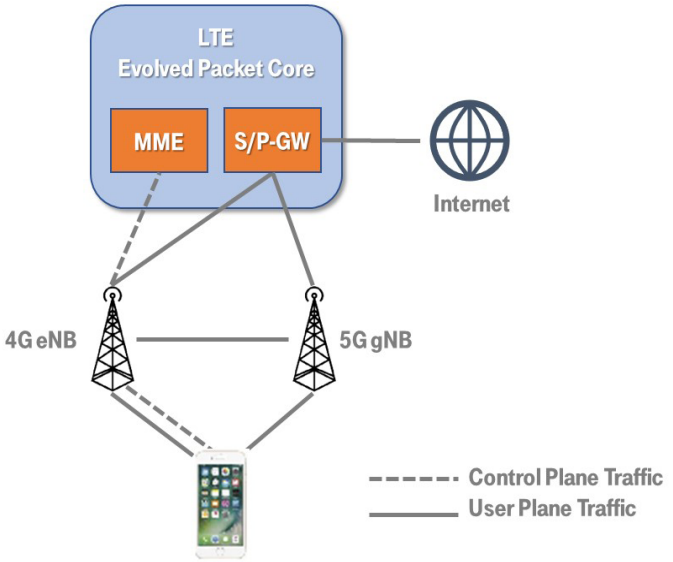


Figure 4: 5G Non-Standalone Architecture

Due to the design of 3GPP specifications for LTE and 5G NSA, these networks are vulnerable to the attacks described in the following sub-sections. This list provides several examples and is not all-inclusive.

1.3.1 2G/3G Downgrade Attack

Downgrade attacks allow for adversaries to force an LTE connected UE to 2G or 3G, which has significantly less security controls. Ultimately, adversaries could perform man-in-the-middle (MiTM) active attacks and/or a passive (e.g. eavesdropping) attacks to collect sensitive information. A customer experiencing abnormal behavior in their LTE connection could indicate of this type of attack. For

example, if customers have historically experienced reliable 4G/LTE connections in a specific, well known area (e.g. work, home, coffee shop, etc), but their connections suddenly revert to 2G/3G . Customers can typically tell if they are connected to 2G or 3G from the connection indication on their phones. Instead of indicating LTE and/or 4G, the screen will typically display an "E", "G", or another symbol.

1.3.2 IMSI Tracking (Privacy)

The IMSI (International Mobile Subscriber Identity) is a unique number that can be captured in the clear over-the-air. High cost Stingrays are no longer required for this attack, because low cost software defined radios (SDRs) can be purchased over the Internet. This could allow bad actors to pursue lower value targets resulting in privacy concerns for the general public. These same, low-cost SDRs would more likely be used by an adversary to track and exploit higher-value targets for various reasons. Adversaries could determine the value of the target based upon the movement of that target. For instance, if the target regularly visits a defense industry research and development facility, a military facility, and/or an executive office's administration facility. This information could indicate that the target has potentially significant sensitive national security related information. Once a high-value target is identified, additional tracking could expose if the target is engaging in inappropriate personal activities that could be used to compromise the individual's integrity.

1.3.3 Man-in-the-Middle Attacks

The Access Stratum (AS) over-the-air User Plane traffic is not adequately protected by Integrity Protection security algorithms. This potentially translates to a scenario where a customer's message and/or communication flow could be intercepted in the middle between the UE and the server. An adversary could manipulate the customer's message and/or communication flow between the UE and the server. If the customer's communication is protected by end-to-end security encryption protocols (e.g. SSL, TLS, IPSec, VPN, etc), then this attack is impossible. Almost all corporate, business, and social media communications (e.g. Corporate VPN, Banking, Facebook, Twitter, etc). are protected by end-to-end security encryption protocols. The threat against the customers and their privacy is limited but worth noting, as not all Internet destined communications are protected by end-to-end security encryption protocols. It is recommended that the customer validates the SSL certificates. These certificates should be valid for their website traffic, and cautious of any websites where the SSL certificates may be expired, self-signed or have questionable domain names. Customers should verify that websites and/or business are validated through Extended Validation [5] (EV), which provides the highest level of site security.

1.3.4 LTE Roaming

LTE roaming is heavily dependent upon the SS7 and Diameter protocols [6]. SS7 was initially developed in 1975 [7] for the publicly switched telephone network (PSTN) but was the foundation of voice communications in 2G and 3G networks. Diameter [8] is an authentication and authorization protocol defined in 1988 to supersede the RADIUS protocol. Both the SS7 and Diameter protocols have been used in large scale, and have had known security vulnerabilities that have been the focus of attacks for years [9]. As a transition from SS7, many operators have deployed voice over LTE (VoLTE), which uses Session Initiation Protocol / Real-time Transport Protocol (SIP-RTP) instead of SS7. Diameter is still used in LTE for authentication, authorization and Policy Charging and Control (PCC) functions. Diameter and SS7 are vulnerable to eavesdropping including voice calls, reading text messages, and tracking phones. Some LTE roaming mobile network operators and mobile virtual network operators do not support VoLTE, so even if an operator has deployed VoLTE and its customer roams into an MNO/MVNO network that does not support VoLTE, then home networks must use SS7 for voice services for that roaming customer. Many operators have SS7 and/or Diameter firewalls but these firewalls are subject to a number of cross-protocol attacks [10].

1.4 Threats, Vulnerabilities and Attacks—5G Standalone

Starting in late 2020, operators will begin to deploy their 5G Core Networks which will allow them to deploy 5G Standalone services. In addition, the 5G Core Network along with the implementation of 3GPP Release 16 specifications will allow for the commercialization of the Massive Machine-Type Communications (mMTC) [4] and Ultra-Reliable Low Latency Communications (URLLC) [4] use cases.

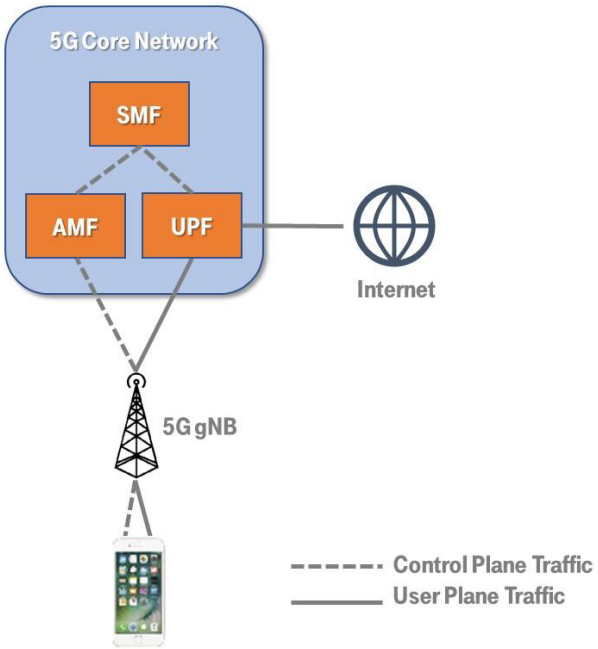


Figure 5: 5G Standalone Architecture

As described in Section 2.1, key aspects that differentiate 5G SA from 5G NSA and older versions of wireless networks are areas of improved wireless privacy that conceal the IMSI, a Service Based Architecture (leveraging Cloud-Native techniques,) SDN and NFV.

1.4.1 Service Based Architecture

The 5G ecosystem is largely composed of software that could run-on general-purpose hardware that communicates with application programming interfaces (APIs). The integrity of the software, especially from open-source locations and the overall software supply chain, is an area of vulnerability. 5G leverages Cloud-Native principles, where services can be created, destroyed, and constantly communicating with each other in a dynamic fashion. All systems must be properly authenticated with protected communication to prevent unauthorized instructions, or the unauthorized access to resources (having the ability to instruct the system to exhaust its resources is one form of DoS attack). In the 5G architecture users will have access to network specific services. Any current hardware or software faults (including operating systems) will also exist in 5G architecture. Ultimately, 5G hopes to use the concepts of SDN and NFV, and both come with unique threats and vulnerabilities.

1.4.2 SDN

At its heart, 5G systems and functions are programmable software modules. Much time and research has been devoted to understanding the security concerns of programmable modules in large systems. The ability to program functionality means the operator, or even the user, can change the way the entire system's or software's behavior. As such, it's crucial that only authorized entities have the ability to change or program the network; the provider vets and controls the end-user's capabilities. This applies to both human-selected and automated service chains. Prior to implementation—and if programmability is available—there must be a way to test if changes in behavior produce the desired result instead of unintended outcomes. Network verification techniques are currently being used to ensure that a network is adhering to its intended policies. It's not clear if the network verification system itself can be vulnerable to attack, which might deceive the operator into believing policies are being adhered to when they are not.

1.4.3 NFV

Both the integrity of the code that comprises a virtualized function and the interaction between virtualized functions themselves is important. Open-source software is a concern in any environment where it can be used. The functions that comprised the 5G system can be composed of open-source software elements, but their security and integrity is not always known. The virtualized elements must communicate with each other in a standardized, API-style environment. The APIs themselves must adhere to standards but must also have safeguards in place to avoid being manipulated in unintended ways to cause disruption.

1.4.4 Supply Chain—Hardware and Software

Hardware supply chain security is a well-known area for most operators and OEMs/ODMs. Software supply chains are mostly unknown, present, and significant security threats for operators and OEMs. Both hardware and software supply chains need to be viewed with a zero-trust approach from a Network Operations centric perspective that is vertically implemented into the supply chain. The operators need to mature their Supplier Risk Management Processes and Third-Party Risk Management Processes to incorporate the new open-source platforms and open-source operating systems, and vendor software deliveries must be scanned and/or routinely updated or hardened. The operators must require that the OEMs embrace a secure and reputable auditor. Compliance reports must be delivered per GSMA's Network Equipment Security Assurance Scheme (NESAS) [11], or an audit from a trusted outside entity, such as the System and Organization Controls for Supply Chain (SOC for SC) from AICPA [12]. Incorporating the NESAS auditing by the OEMs via an independent, reputable, 3rd party auditor will ensure that the OEMs follow best practices towards secure software development and securing the end-to-end supply chain.

1.4.5 Man-in-the-Middle—User Plane Integrity Protection

Like 4G, 5G is plagued with a lack of User Plane (UP) Integrity Protection (IP) for the Access Stratum over-the-air flows. 3GPP defines UP IP integrity algorithms in their specifications but many of the OEMs have not implemented them because of impact on the user experience (e.g. download and upload data throughputs). Enabling UP IP requires considerable compute resources and adds overhead that directly impacts the maximum throughputs that can be measured on the user device. IP is enabled on the Control Plane messages but that still leaves the user's data traffic vulnerable because the Control Plane and User Plane are segregated. For example, the lack of UP IP could enable a rogue base station to manipulate the user data messages (i.e. DNS) and redirect a user to a malicious website. In 5G, the 3GPP SA3 Security working group is reviewing UP IP specifications that encourage OEMs to develop the capabilities in the gNB radios, and user equipment 5G baseband modems. Some 5G radio OEMs are developing hardware-based encryption acceleration to mitigate the negative impacts on the user's download and upload data experience.

1.4.6 SUPI/SUCI Privacy

In LTE, the IMSI is sent "in the clear" over-the-air allowing for an adversary to capture and track a high-value customer. 3GPP has addressed this by removing the clear-text IMSI and moving to a SUPI/SUCI solution. The IMSI in 5G is now called a Subscription Permanent Identifier (SUPI). The Subscription Concealed Identifier (SUCI) is sent over-the-air rather than the SUPI, which prevents an adversary from tracking a high-value target. As of today, there are a few instances where the customer could be tracked if the SUPI is sent over-the-air. Those cases are:

1. Unauthenticated devices attempting to make an emergency call
2. A UE with a legacy SIM that has not been provisioned with the 5G public key via an OTA update
3. Customers bringing their own devices into an operator's network with SIMs that have not been updated

3GPP SA3 is working to restrict use cases (2) and (3) by forcing the SIM OTAs to be done over the LTE network

1.4.7 5G NSA and SA Roaming

5G NSA roaming is essentially 4G roaming because NSA uses the EPC for all Core Network functions. From a security perspective, a 5G NSA roaming connection introduces no new protections since it continues to use Diameter, SIP/VoLTE and possibly SS7. Diameter and SS7 are vulnerable to eavesdropping including voice calls, reading text messages, and tracking phones.

For 5G SA and to mitigate these vulnerabilities, the roaming flows will be considerably different, as HTTP/2 and JavaScript Object Notation (JSON) will be used versus the legacy Diameter protocol. Later in 2020 and/or early 2021, operators will

start deploying Voice over New Radio (VoNR) which will replace VoLTE in the 5G network. 5G introduces a Security Edge Protection Proxy (SEPP) which will establish a secure, encrypted connection with the roaming partner's SEPP. The SEPPs will then pass HTTP/2 SBA control plane flows for authentication and authorization. For the data flows, there are two options available:

1. Securely between the VPLMN (Visited Public Land Mobile Network) and the HPLMN (Home Public Land Mobile Network) UPFs (aka Home Routed)
2. Through the VPLMN's UPF (aka Local Breakout)

The Control Plane messages will be passed securely between the VPLMN and HPLMN Security Edge Protection Proxies (SEPPs) via the N32 interface. The User Plane traffic is carried over GTP-U. For Option (1), GTP-U does not provide confidentiality or integrity protection of the user's data traffic, so a secured tunnel between VPLMN and HPLMN should be used to protect the user data flows on the N9 interface.

1.4.8 Interoperability for 5G SMS over NAS

A 5GC network could be exposed to 3G/4G vulnerabilities when a 5G network operator interoperates with their legacy 3G/4G systems, or interconnects through an inter-carrier IPX provider. 5G Short Message Service (SMS) over Non-Access Stratum (NAS) is a clear example of how 5GC networks could be exposed to SS7 or Diameter vulnerabilities.

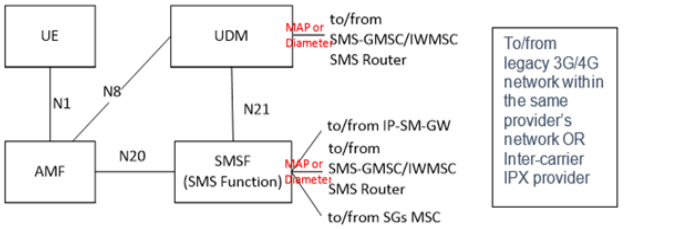


Figure 6: System Architecture for 5G SMS over NAS

Figure 6 illustrates two areas of potential concern: SMSF and UDM interoperability with legacy 3G/4G networks.

The 5G AMF provides the path for Mobile Originated (MO) or Mobile Terminated (MT) SMS for 5G subscriber over the encrypted NAS protocol on the N1 interface, like the MME in 3G/4G networks. The SMSF (SMS Function) is like the Mobile Switching Center / Visitor Location Register (MSC/VLR) in 3G/4G Core Networks. It relays MO and MT SMS messages between a 5G subscriber and legacy 3G/4G networks. An attack surface potentially opens at the SMSF to Short Message Service Center / Gateway (SMSC/SMSGW) interface with the SS7 MAP protocol specified in 3GPP TS 29.002 or the optional Diameter reference point SGd specified in 3GPP TS 29.338.

The Unified Data Management (UDM) node is like the Home Location Register / Home Subscriber Register (HLR/HSS) in a 4G network. The legacy 3G/4G SMSC or SMSGW connects to the UDM to get a 5G subscriber's connected state and determines whether to deliver SMS to the 5G subscriber or not. This connection potentially opens another attack surface because the UDM to SMSC/SMSGW interface is based on either the SS7 MAP protocol specified in 3GPP TS 29.002 or Diameter reference point SGd specified in 3GPP TS 29.338.

2. Security Considerations & its Impact on 5G

2. Security Considerations and its impact on 5G

2.1 Cloud-Native Architecture and Security

As the Core and Access of 5G shift to Cloud-Native architectures, several issues must be considered:

- Should one build their own cloud infrastructure or leverage a pre-existing one (i.e., buy versus build)?
- How does one take advantage of high availability and resiliency within a specific cloud?
- How can communications between processes—within the same cloud or across clouds—be secure?

2.1.1 Cloud Native MEC

Multi-Access Edge Computing (MEC) will deliver 5G services from a distributed cloud enabling new mobile use cases for consumers, business, and public safety. MEC is a distributed cloud with multiple sites that bring the execution environment geographically closer to the UE. The main benefits of MEC are low latency, high bandwidth, device density, data offload, and trusted computing and storage. These benefits enable 5G use cases like enterprise private networks for industrial IoT, smart cities, autonomous vehicles, remote surgery, virtual reality, cloud gaming, and high-quality video streaming without buffering.

MEC offerings will be built upon microservices within SBA in which network operators evolve from their NFV infrastructure (NFVI) architectures using VNFs to a cloud-native infrastructure using Containerized Network Functions (CNFs). This cloud-native approach provides flexible, automated, and scalable grouping of microservices-based applications in separate containers. Continuous Integration/Continuous Delivery (CI/CD) with Life Cycle Management (LCM) of microservices increases service velocity while improving resiliency. LCM is implemented through dynamic orchestration and management of workload services on top of the distributed cloud infrastructure. The Cloud Native Computing Foundation (CNCF) has certified Kubernetes for orchestration of microservices running in Docker containers in a distributed Containers-as-a-Service (CaaS) infrastructure.

Cloud native MEC provides inherent security protection:

- Each container performs a dedicated function, enhancing behavior profiling and anomaly detection.
- Isolation of containers prevents spread of malware and viruses.
- Decomposed software provides efficient software version updates and security patches.
- Compute services in cloud-native applications are designed to be ephemeral, reducing the attack surface.
- Resiliency is gained with increased speed to start a container and horizontal scale to dynamically respond to threats.
- Segmentation with network slicing separates traffic and isolates compute resources.

However, cloud native MEC architecture also introduces new security risks:

- The uses of open source code, more interfaces, and new APIs introduce new threat vectors
- Shared hardware resources can result in cross-contamination
- Vulnerabilities in the shared host platform, Container-as-a-Service (CaaS) and Platform-as-a-Service (PaaS) can impact the container security
- Containers requiring elevated privileges can cause security risk to both host as well as other tenant containers
- Dependency upon central orchestration introduces a new threat vector
- High data volume and sessions increase risk from an attack
- Applications running in a microservice architecture are as vulnerable to the same attacks as traditional applications

There are several steps that can be taken to mitigate risk:

- Application developers need to practice Security by Design utilizing the "Shift Left" and "Shift Up" design concepts. Shift Left is a development practice to prevent defects early in the software development process. Shift Up is the practice of ensuring flawless execution of applications inside containers.
- The cloud native architecture must implement Zero-Trust with secure authentication and authorization on the control plane.
- IPSec and DTLS should be used for control plane and data plane protection (as specified in 3GPP TS 33.501, 5G; Security architecture and procedures for 5G System).
- Perimeter protection can be implemented with DDoS Protection, Behavior Profiling, and Anomaly Detection.
- Threat Intelligence can be leveraged for real-time detection and mitigation of malicious attacks.

MEC leverages a combination of different sizes of centrally orchestrated and managed cloud data centers distributed at global, national, regional, and local locations. MEC infrastructure can be managed or hosted by Communication Service Providers (CSP) or Hyperscale Cloud Providers (HCP), such as AWS, Microsoft Azure, and Google Cloud. CSPs can choose different roles, a single role or a combination of roles depending on their ambition to build distributed edge infrastructure.

Cloud-native MEC can be implemented as a Private Cloud, Public Cloud and Hybrid Cloud:

- In the Private Cloud model, the CSP hosts its telco and non-telco workloads in its own private data centers. This uses the Infrastructure-as-a-Service (IaaS) model.
- In the Public Cloud model, the CSP deploys its non-telco workloads in the HCP's public cloud. The HCP provides distributed infrastructure and the cloud-native platform. The HCP orchestrates and manages the CSP's workloads. This uses the CaaS model.
- In the Hybrid Cloud model, the HCP provides distributed infrastructure to the CSP. The CSP builds, orchestrates, and manages its workloads on the HCP infrastructure.

The advantage of the Public Cloud and Hybrid Cloud models for the CSPs is that MEC can leverage the HCP's distributed data centers for rapid build-out. However, the Public Cloud and Hybrid Cloud models cause a shift in the threat landscape increasing security risk because a third party HCP is introduced into the ecosystem by providing hardware and software APIs. The CSP owns the workloads but not the infrastructure, causing the CSP's security team to effectively cede control to the HCP. The CSP and HCP must collaborate to build an effective security posture that clearly defines:

- secure communication across domains
- security perimeters
- demarcation of enforcement points to automatically detect and mitigate attacks and breaches
- security management with defined roles and responsibilities to properly respond to security events; and
- APIs for communication and data exchange across domains that are built with strong security including identity management, access control, Public Key Infrastructure (PKI) authentication, anomaly detection, and logging.

2.1.2 Risks of Open-Source in 5G

Use of open-source will continue to increase as 5G operators and vendors rely on open-source software to speed delivery of new solutions and reduce total cost of ownership. Open-source software can be considered analogous to corporations outsourcing functions not related to their core competencies. This introduces a new set of security challenges to keep a consistent and coherent assurance of security-by-design and prevent resulting security flaws.

While the use of open source offers benefits of time to market, cost and reliability, it also can be the source of vulnerabilities that pose significant risk to application security. Many 5G vendors and operators rely on open-source software to accelerate delivery of digital innovation. Both traditional and agile development processes frequently incorporate the use of prebuilt, reusable open-source software components. As a result, some organizations may not have accurate inventories of open-source software dependencies used by their different applications, or a process to receive and manage notifications concerning discovered vulnerabilities or available patches from the community supporting the open-source.

Open-source software provides attackers with a target-rich environment because of its widespread use. Open-source software is incorporated into applications in many ways, and oftentimes, 5G operators or vendors will not know where open-source is used. When open source is used as the foundation for a vendor's product, any vulnerabilities could threaten the integrity of the vendor's solution.

Asking vendors to disclose the open-source components used in their products may disclose more vulnerabilities and add to the risk. Note that there may be several vulnerabilities in a specific software component, but they may only exist as stand-alone components that can be nullified when incorporated into the vendor's solution. For example, through middleware where the open-source component is isolated. Care must be given in how open-source components are disclosed to prevent exposure.

2.2 Software Driven Operations

As 5G leverages aspects of the Cloud and software control, service provider's operations need to change as a result. Operations is a collection of telemetry, thresholds, triggers, and processes that dictate how to react to a changing network environment. To have software driven operations, each step in the workflow needs to be fully understood. All processes must be fully documented so that each step can be turned into properly functioning software code. Once created, these detailed workflows must be tested, implemented and monitored.

All routines and APIs require strong authentication. Separate roles and authorizations are required for each routing and API. It is not sufficient to authorize at the system or person level, as lateral movement in a compromised software driven operation can be devasting. Control to the routine and code level is mandatory.

A response team needs to have coders who can dissect what might have happened during an incident. Did the code execute as intended, or was the code poorly written, causing a process to hang or malfunction? All software routines should be logged and audited. It's not enough to have operators produce post-event incident reports; there should be a way to verify precisely what happened and when by analyzing the logs.

2.2.1 Automation

Automation is the process of turning manual human processes into independently running machine processes. At a high level, this requires the collection of data (telemetry), decisions on what actions to take (or not), and then the implementation (action) of those decisions. This process runs in a continuous loop, either with human interaction or not, and continuously refines the behavior of a system towards its desired state.

Gates of trust are needed to properly implement an automated system in a 5G network. Some actions might be easy to automate early on, whereas other actions need to rely on human interaction. Sufficient auditing must be in place to verify that the series of automated events achieve the intended outcome.

2.2.2 Orchestration

Orchestration is the concept of managing the interaction between the different processes. In 5G, both the RAN and 5GC networks will be based on SDN, NFV and cloud-native infrastructure. Within a specific network, orchestration will help coordinate the use of many different 5G resources. One example is the creation and management of network slices: the slice is created with the correct performance characteristics and is connected to the correct set of UEs. Like general software processes and automation previously described, orchestration must have secure communication between

processes, and have the correct authorization to control resources to operate properly.

5G networks interoperate, and 5G has important security improvements for roaming with the introduction of the SEPP in the 5GC. However, the operations groups between various network providers must also interact. Here questions arise regarding the tier structure of each operations group, and what, if any, shared control might need to exist between operators. If so, what are the requirements for federated identity, control and limited access and authorization between operator's infrastructures? These are among the question yet to be answered.

2.2.3 AI/ML

Artificial Intelligence (AI) and Machine Learning (ML) represent a hands-off approach to dealing with new scenarios that automation and orchestration haven't anticipated yet. One issue is that AI/ML is often "Blackbox" and is not fully understood by the operator. The algorithm that was produced by AI/ML was learned or trained on a data set, but understanding the key influencers were in the data set to drive the specific actions is crucial. Advanced forms of AI will need to add context, and necessitates the ability not only to understand what happened, but also why. This is necessary so bad actors do not acquire the ability to influence the AI and manipulate the outcomes for their own purposes.

2.3 NFV Security

Operator's networks have traditionally leveraged OEM proprietary hardware that is purpose-built for specific functions. That has made it very difficult for the operator to scale the networks, particularly during peak demands like natural disasters. Some operators started retrofitting some of their LTE core network functions moving away from these propriety hardware platforms and replacing them with open-source platforms (e.g. OpenStack, Red Hat Linux, etc.) running OEM network function (NF) specific applications (NFV) to prepare for 5G NSA. The 5GC network may become completely virtualized using ETSI's NFV Architecture [13] which allows for operators to deploy scalable, elastic, and highly reliable networks.

Security considerations exist at the virtualization layer and lower supporting layers. This includes the virtualization infrastructure (the NFVI, which holds virtual compute, networking, and storage resources), as well as software on compute nodes such as hypervisors, host operating systems, and container run-time systems. Management and control of the elements mentioned above include the virtual infrastructure manager (e.g. OpenStack) and container orchestration engines (e.g. Kubernetes) which support the full lifecycle management of VNFs and CNFs that carry out the network functions.

NFV security considerations apply across the entire virtualization infrastructure and its management plane. Security considerations specific to the particular functionalities of a Network Function is described in other sections of this paper.

2.3.1 Management & NFV Orchestration Security

In the event of a natural disaster, an operator may spin up new regional and/or market level NF virtual machines and/or containers to provide additional network capacity delivering an elastic network based upon demand. NFV Management and Orchestration operations must be made secure, especially the life cycle management of VNF/CNF workloads.

Operators must ensure that they use an out-of-band (OOB) management network that is not accessible from the Internet so management interfaces are secured from remote access. If an operator allows for remote access into the OOB for employee and/or OEM support, the operator must ensure that they use a multi-factor authentication (MFA), at a minimum, for any type of VPN access. An MFA VPN combined with Zero-Trust greatly improves secure remote access to protect the 5GC Network.

2.3.2 Resource Isolation and Securing Traffic

NFV infrastructure shares resources among multiple tenants and services. This may require isolation between tenant software workloads as well as separation of traffic at both tenant and service levels. Rules for the placement of instances of software workloads in the infrastructure will

depend on the level of isolation required. An NFV system can stretch across multiple infrastructures, involving multiple tenants, which may benefit from the creation of multiple trust zones even within a single service provider network.

Traffic routed through a virtualized network may not be completely accessible for physical firewall controls or visible to traditional security inspections as previously applied on physical networks. This raises the need for virtual security appliances like virtual firewalls or virtual IDS/IPS to achieve a level of security comparable to traditional networks.

For public facing interfaces, operators need to effectively monitor the NFV platforms from denial-of-service (DoS) attacks, security anomalies, and known security vulnerability attacks. Operators need to use all the layers of security (defense in depth) to protect the NFV platforms including firewalls, access control lists, IP tables, rate limiting, closing all unnecessary ports, disabling all unnecessary services, using strong confidential integrity algorithms, and so forth.

2.3.3 Software Centric Security

Numerous interworking software artifacts that make up an NFV system that demands new software centric security considerations:

- New NFV platforms operating on open-source software provide a new attack surface with known vulnerabilities, whereas the OEM proprietary hardware platforms—most still Linux based—were slightly more protected because the vulnerabilities were not always made public
- Tamper proof records of all installed and/or running software, configurations and versions must be maintained.
- Virtual and physical resources each require a current mapping (Which VNFs are running on which physical resources?).
- The ability to prohibit or disable sharing of memory between workloads and to restrict workloads from accessing special memory locations is necessary.
- Attestation of Software Components

Operators need to define, implement, audit and institute a governance team to confirm that the NFVI platforms (i.e. OpenStack, OS, etc) and the VNFs are patched regularly—particularly when Coordinated Vulnerability Disclosures (CVDs) are issued and/or mitigated.

Typically, VNFs/CNFs are distributed from suppliers to operators in software packages which go through an ingestion/onboarding process. As NFV systems mature, this supplier-operator pipeline will become increasingly automated, which means the authenticity and integrity of the VNF/CNF packages must be guaranteed. In addition, there is a need to authorize the onboarding or instantiation of each individual package in the proper target network location (i.e., to avoid a package intended for one location to be loaded or spun up in another location). Finally, software centric security frequently needs confidentiality of software package contents comprised of artifacts like executable code, scripts, and network configuration after onboarding into an operator's repository.

2.3.4 NFV Security Monitoring

The NFV paradigm which is based on virtual infrastructure has different security interface characteristics and requirements for virtualized tests and monitoring functions. Traditional techniques of monitoring use standardized monitoring interfaces on physical nodes, active/passive hardware probes, deep packet inspection, and correlation with control plane and management plane information. However, they fall short in an NFV environment due to much less overall visibility. This includes the concealment of monitoring interfaces within shared memory and virtual sockets, the hidden data flows within a single VNF (i.e., between VNF components), between VNFs on same physical host, and flows within virtual switches and routers.

Obtaining a complete security picture also requires knowledge of orchestration activities, as well as view across service chains, and integration of security mechanisms provided by the infrastructure (NFVI) and workload (VNFs/CNFs) domains. This calls for the potential need for:

- active/passive/hybrid monitoring using a combination of agents and agent-less approaches (e.g., virtual probes)
- monitoring the management plane as well as the service plane (workloads)
- securely monitoring logs system wide
- securely monitoring resource usage

2.3.5 Attestation of Hardware/ Software Components

NFV systems being composed of predominantly software components face threats that allow attackers to modify the software or inject new software that elicits malicious behavior. In other words, the execution of code that was not intended must be detected. As a result, attestation is used to determine the trustworthiness of the NFV system components.

Attestation needs to cover all the NFV components: hardware platform, virtualization layer and infrastructure components, as well as the VNF/ CNF Software (workload). Typically, attestation must be anchored on a hardware root of trust. It is also commonly referenced as Remote Attestation, in which an independent party remotely conducts verification of the integrity of the NFV components.

2.4 SDN Security

SDN is characterized by the separation of an application plane, control plane and a forwarding plane (as defined by the Open Networking Foundation [14]). 5G utilizes two concepts to achieve a similar architecture: the SBA for the application plane separation, and CUPS for the control plane and forwarding plane separation.

There are several aspects to SDN security:

Logically Centralized Control: SDN uses the concept of a logically centralized controller (which may be physically distributed for resiliency). This then creates a location where, if compromised, an attacker would have control over the entire system. The programmability of the system is tied to the business management systems (OSS/BSS), and any real-time control of one entity should be fully isolated from that of all others.

APIs and Protocols: Communication between each plane of the SDN architecture is achieved via APIs and/or SDN-specific protocols through various virtual interfaces. These interfaces can be internal or external, which introduces new attack surfaces as that communication can be spoofed, disrupted, or used to deny service. It is therefore critical that the appropriate level of authentication and authorization is applied to all communication interfaces. Additionally, the packets themselves must have ensured integrity, confidentiality, and rate control. All of this must be extended to a federated environment, where the controllers or elements of one provider must communicate with elements in another provider.

API Security: API security involves securing data end-to-end, which includes security from a request originating from one network element to another, passing through other network elements. API security can include a) data in transit security (securing data in motion between the control plane, user plane and services,) and b) access control and security against Denial of Service (DoS) Attacks. For instance, the 5G network elements can leverage shared keys, values generated by some type of encryption algorithm. The keys can be kept in a central secured system. During the communication process, the central system will validate whether the shared key in its records to authenticate the communication. In addition, policy enforcement of the APIs (e.g., rate limit, discard, allow) controls and ensures only the authenticated 5G network elements can exchange communication.

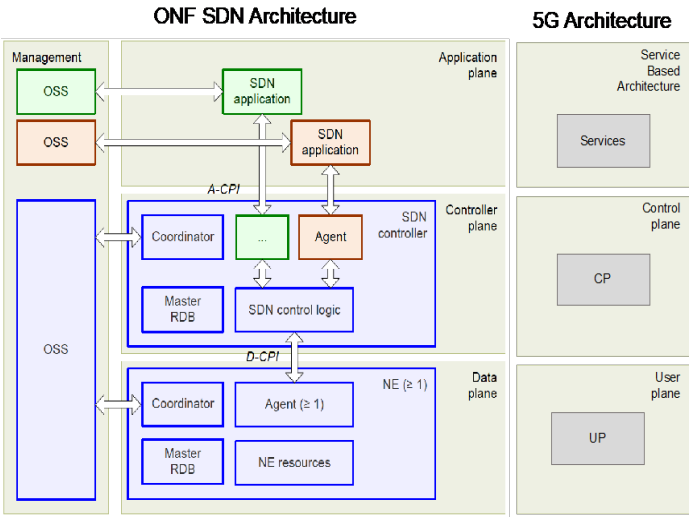


Figure 7: SDN Architecture and 5G

Other Considerations: In addition to the points above, a robust SDN architecture should include:

- **Scalability:** all elements of the architecture should be able to scale up and scale out.
- **Extensibility:** as technology and innovation progresses, the architecture should have the ability to be updated or augmented with new capabilities.
- **Backwards compatibility:** like previous 3GPP wireless generations, multiple iterations of technology will need to interoperate with previous ones.
- **Reduction of complexity:** every system should be designed with the minimal amount of complexity, as increased complexity leads to increased costs and can compromise reliability, and ultimately security.
- **Support for automation:** networks above a certain size, become increasingly difficult, if not impossible, for a human to manage. Therefore, automation is necessary to achieve a proper scale.
- **Ability to monitor, troubleshoot and debug:** Software drive operations rely on the ability to monitor, audit, troubleshoot and debug (as automated as possible) any issues with the system.

2.5 Data Security at the Edge

In the edge and far edge, operators are introducing virtualized network elements (NEs) based on Container technology which causes a new set of security challenges. Containers may require elevated privileges to support certain network functions that could cause security vulnerabilities for the host system as well as peer containers. Careful design and deployment considerations should be allocated to address such concerns. Efforts should be made to ensure that no aspects of these systems are compromised.

Edge NFVIs need to be secured physically. From a software point of view, the access control for the entire computing platform and NFVI components must be locked down. Security audits are warranted for an edge boilerplate configuration in addition to core implementations.

In many Edge deployments the transfer and caching of subscriber and organization affiliation credentials are necessary. An example of this might be a PLMN's private key. Edge Network Elements used for legal intercept should be deployed with the same, if not more stringent, security policies as in traditional core networks.

Table 1 consists of issues and recommendations were developed by the Next Generation Mobile Network (NGMN) Alliance in the paper, "5G security – Package 3: Mobile Edge Computing/ Low Latency / Consistent User Experience version 2" [15]. The issue and recommendations are still valid today.

Table 1: 5G Issues, Recommendations and Standards Status

Potential Issue	Recommendation	5G Standards Status
1) Billing Data Manipulation	System and Interface Hardening. UE cross-checking. Integrity Protection.	All recommendations are available and addressed.
2) 3rd Party AFs running on same platforms	VM and Container Security 3rd Party Audits Legal Indemnification	Proper NFV separation available. Audit process specified – can apply to AFs as well. Legal Indemnification can be handled a business process.
3) Third party applications allowed to influence the network	Application Policing	Sections 3.1-3.3 along with Slicing provide for policing and resource segregation.
4) Services to third party MEC application providers	Mobile operators supply security assurance services to 3rd party apps	The audit process, signed code similar to application assurances on UEs.
5) User Plane Attacks in a Mobile Edge Computing Environment (local DNS and cache DDoS Attacks)	Duplicate Caches and provide anti-DDoS, certificate management combined with DNS. Caches change as user moves.	All of these are available. Future Inclusion of user plane integrity would also mitigate these attacks. Cache change based upon mobility can occur by triggers and application-specific signaling.
6) Storage of Sensitive Security Assets at the Edge	Encryption of sensitive data bad protection of keys.	Most general computing platforms contain trusted environment facilities with tamper proof credentials for encryption of sensitive data.
7) Exchange of Sensitive Security Assets between core and Mobile Edge. Man in the middle threats to privacy, etc..	Encrypt this information	2-way TLS Authentication, Encryption and Oauth2 Authorization are mandated in this scenario. NDS with mutual authentication is alternative.
8) Trust Establishment between functions at the core and at the edge	Mutual Authentication on process start/restart. Network based Code signing verification as part of Authentication	While the PKI management and other forms of credential management are outside scope of 3GPP the other requirements mandates it. Standards indicate that Mutual Authentication MUST be done.
9) Security of Communications with the MEC Orchestrator	Same as above.	MEC and other Edge platform initiatives all have this in place.
10) Law enforcement requirements for MEC deployments	All of the above. Physical/Access controlled environment. Periodic Audits as well.	These functions are available and are secured. Physical Access Control understood.

Other remediations, mitigations and recommendations in this paper are also applicable to Edge Computing. Please refer to Section 4.5 (Slicing) for additional recommendations.

2.5.1 Edge Computing Frameworks for Orchestration and Business Logic

Many competing software frameworks have been developed to support NFV and SDN enabling Edge Computing.

The 5G America's whitepaper, "5G at the Edge" [16], delves into many of the frameworks. Each framework accounts for separate ownership of computing resources with an optional split in orchestration (localized and core) and business logic. The interfaces between these orchestration elements should be secured with mutual authentication and integrity protection.

In terms of network functions that have legal intercept requirements (e.g. user plane functions (UPF)), such ownership models and even physical placement in common access-controlled areas are likely not feasible but other functions such as applications are.

2.6 URLLC and User Plane Security Considerations

UE power budget, latency and security of user plane traffic are often adversarial goals in application solutions. Most general-purpose computing platforms provide for line speed encryption and integrity protection; however, they typically increase power usage on both the UEs and network functions.

Tethered and Untethered machines, such as in a factory, are expected to have adequate power budgets, and possibly employ wireless charging to eliminate charging downtime. V2X, agriculture and mining solutions provide adequate power budget as well.

Security of user plane traffic is also a reliability attribute. Enabling integrity protection on the user plane can thwart denial of service attacks. Applications could employ such security end-to-end to avoid enabling encryption and integrity protection between the UE and the Core. These solutions often employ tunneling solutions that can create MTU issues and additional latency. This method is not recommended for very low latency communications, as enabling the security in the core is probably a better alternative.

2.7 Private 5G Networks

The term "private networks" has been around for a while but has evolved over time. The new abilities of 5G in terms of broadband speeds, low latency and reliability open many new solution-oriented use-cases for private networks. Below is a list of solutions currently in use today for private networks:

- **Multi-Operator Core Networks (MOCN):** RAN Sharing with more than one core system
- **Multi-Operator Radio Access Networks (MORAN):** RAN sharing by Bandwidth Splitting
- **Gateway Core Networks (GWCN):** Gateway Core proxies to multiple core systems
- Defining a specific PLMN for an enterprise customer
- Defining an APN or DNN for an enterprise customer
- Closed Subscriber List (CSG) with ACL for enterprise
- Support for 2nd stage authentication of enterprise users

In addition to the existing solutions above, 3GPP has defined additional private network solutions in release 16:

- Formalization of the Non-Public Network (NPN) concept (also known as a Private Network)
- Formalization of Stand-alone NPN (SNPN) defined in Release 15
- Creation of Network Identifier (NID) (NPN ID). Like DNN/APN but unlimited scale for PNI-NPNs. The NID is based upon the NAI type
- Defining a NID, set of slices for an enterprise customer with Closed Access Groups (CAG)
- Tracking Area (TA)-based Geo-fenced Slices for PNI-NPNs
- Support for Per-Slice 2nd stage authentication of NPN users

Many of these solutions were implemented by utilizing the operator's core and placing only RAN equipment in the enterprise sites. Second phases followed the CUPS entrance where User Plane NFs were placed in or near the enterprise. Today compact core networks can be provided in a small rack running on general

purpose compute machines and so the financial barrier has been lowered. Small cells and NFV have also greatly reduced the barrier.

3GPP release 15 began to formalize the standards for private networks with the introduction of Non-Public Networks (NPNs). This first release supported a solution that was a completely disconnected from the MNO, termed Standalone NPNs (SNPNs). They used Extensible Authentication Protocol Transport Layer Security (EAP-TLS) and certificates in many cases instead of IMSIs for authentication of the subscribers.

As these SNPNs were deployed, they began utilizing Non-3GPP access interworking functions (W3-IWFs, e.g., Wi-Fi) to provider's core networks at the sites. This solution is formalized in 3GPP release 16.

Additionally, 3GPP Release 16 added support for NID-based Public Network Integrated (PNI) NPNs and the introduction of Closed Access Groups (CAGs) similar to CSG functions in UMTS and LTE. They also detail specifics around Emergency Services and support the private network based on DNN/APN. MOCN, MORAN and GWCN methods are also recognized in the 3GPP release 16 standards as methods of RAN sharing for PNI-NPNs. Enterprise AAA to each slice using EAP methods is also supported for NIDs.

5G with high bandwidth URLLC support has renewed the interest in private networks. To achieve extremely low latency URLLC many of the core components need to be in or as close to the enterprise as possible. From a security standpoint, the enterprises want their data kept onsite.

Many private networks trials exposed real-world needs which have been folded into the requirements of 3GPP release 17. These needs range from RAN scheduler changes to how to temporarily give access to guest users, similar to how enterprise Wi-Fi user management functions today.

The following scenarios describe security topics of interest for NPNs:

- **5G SNPN:** The security of a 5G SNPN is very similar to a core or Edge deployment. RAN and NAS encryption, and integrity protection should be enabled with NRF equipment stored in an access controlled physical environment. The administrator can create and distribute the NID root pubic key to the UEs.
- **5G PNI-NPN NID and Slices Method:** The security for the slice based PNI-NPN method is handled in the Slicing (section 4.5) and Edge (section 4.6) sections of this document. TLS or NDS (Network Domain Security) is leveraged. Care should be taken in not disclosing a subscriber's affiliation with a particular organization when transferring NID and Slice identifiers over interfaces.
- **5G MOCN and MORAN RAN Sharing:** The enterprise cores are at the location of the RAN deployments, and the cores of the operators are typically at the Edge or deep within the operator networks. The RAN components are recommended to communicate with the operator cores using NDS/IPSEC. Note: there can be many operators—not just one. For example, public safety could be an operator and all service providers could have access in a neutral host environment.
- **GWCN RAN Sharing Method:** A GWCN provides a single logical core reference point to RAN equipment. For 5G this might entail a shared AMF, SMF, and potentially UPF. These would then interwork with other NFs contained in each of the operator's networks. Resource sharing policies could be managed by one of the providers or the enterprise, much like local policy, is defined and implemented for roaming users.
- **Per NID/Slice 2-stage Authentication:** Each EAP hop should be mutually authenticated and provide confidentiality and integrity checks.

2.8 Radio/RAN for New Radio—Jamming, Spoofing and Interception

5G New Radio (NR) is the global standard for a unified, more capable 5G wireless air interface. It will deliver significantly faster and more responsive mobile broadband experiences and extend mobile technology to connect and redefine a multitude of new industries. The 5G NR architecture is compared to the 4G architecture in Figure 8.

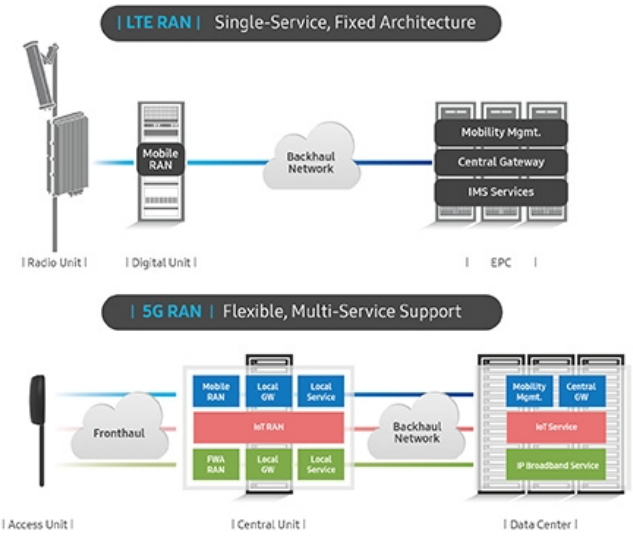


Figure 8: Comparison of 4G and 5G NR Architecture

There are three ways to approach building the RAN in 5G. First, is the traditional method where radio systems (including baseband units) are purchased from the OEMs, like in legacy cellular systems. The second method is to leverage virtualized RAN (vRAN) where centralized and/or distributed baseband units will be connected to remote radio heads using protocols such as CPRI, eCPRI or less commonly OBSAI. The first two options will leverage standardized 3GPP architectures and interfaces. The third method is to leverage open architectures that enable operators to deploy virtualized radio units built upon disaggregated components using Open RAN (O-RAN). O-RAN is being worked within the O-RAN Alliance and will provide operators with an additional option (compared to legacy systems) to meet the growing demand for wireless broadband services. At the time of this publication, the O-RAN Alliance has recently formed a Security Task Force within Working Group 1 (WG1) to address inherent security risks. O-RAN is considered the least secure out of the three options.

2.8.1 Security in 5G vRAN and O-RAN

vRAN and O-RAN provide open, interoperable, and disaggregated virtual networks, as shown in Figure 9 below. Software and hardware components are decomposed, and virtualization techniques are applied to achieve the goals of rapid innovation, technology agility, and integration of best-of-breed components. Two important factors drive improved resilience: improved modularity and reduced interdependencies. Open, interoperable interfaces available deeper within the RAN infrastructure introduce capabilities for isolating controls, greater observability, and independently generated operational telemetry. Those interfaces provide modularity, which could potentially allow more granular security attestation as standards and best practices continue to evolve in this space. They can also reduce dependencies on unique software capabilities, making it less risky to update software to apply fixes.

Given the new nature of applying an open and interoperable approach to the RAN, the real benefits are just starting to be realized in terms of security, agility and overall operational efficiency. vRAN is based upon 3GPP's security standards. O-RAN is built upon 3GPP standards, but the architectural changes and additional interfaces may introduce new security risks specific to O-RAN deployments that must be considered by the O-RAN Alliance. Some examples of O-RAN security risks include, but are not limited to:

- expanded Threat Surface with more functions and more interfacesadditional management interfaces (O2, E2)
  - lower Layer Split (LLS) / O-RU <-> O-DU / O-RAN eCPRI
  - near-RT RIC conflicts with gNodeB
  - near-RT RIC xApps can conflict
- management interfaces that may not be secured according to industry best practices
- increased exposure to public exploits due to use of open-source code

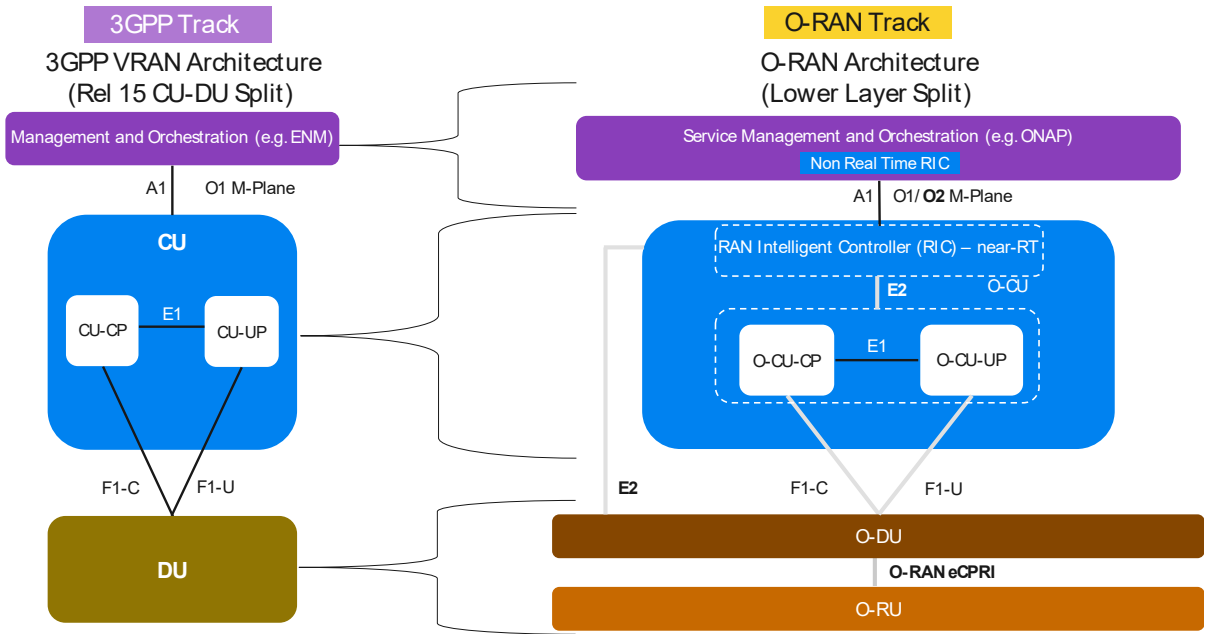


Figure 9: Efficiencies and Threats for an Open 5G RAN Architecture [17]

2.8.2 Updated RAN Security in 5G

This section describes the main vulnerabilities and threats associated with the 5G RAN. In recent years, a large body of literature has revealed numerous security and privacy issues in 4G mobile networks. Most of the published attacks at the 4G RAN layer involve Rogue Base Stations (RBSs) or IMSI catchers to target individual user(s) during the UE's initial attach procedure to the network or paging attacks using the IMSI paging feature. In such attacks, the information obtained on IMSIs may be used later for other types of attacks. Fortunately, 5G technology and standards are addressing these known IMSI threats. 5G utilizes a Subscription Permanent Identifier (SUPI) which is often referred to as an IMSI. Instead of sending the SUPI over the air like 4G sends the IMSI in the clear, 5G will encrypt the SUPI and send the encrypted SUPI over the air (aka Subscription Concealed Identifier—SUCI). In 5G, the SUCI will be sent in normal situations but there are a few scenarios where 3GPP has approved the transmission of a null encrypted SUCI (aka SUPI). For example, an unauthenticated device attempting to use emergency services (e.g. 911) will send the SUPI over the air. An Operator may allow for the SUPIs to be used versus SUCIs which does not improve security related to IMSI tracking.

5G RAN deployments may use massive Multiple-Input Multiple-Output (MIMO) antenna arrays and beamforming to deliver higher capacity and improve the customer experience. In addition to existing low band and mid-band spectrums, some operators may use millimeter wave (mmWave) spectrum. It is not expected that mmWave by itself is less secure than any other frequencies, but the prevalence of off-the-shelf radio sniffing and/or rogue base stations are going to be more difficult to source. The data and signaling transmitted and received at the radio layer is expected to be appropriately encrypted with protected integrity per 3GPP specifications, and at higher user data layers whenever possible (i.e. SSL, VPN, etc.).

Rogue Base Station (RBS) threats will persist in 5G deployments as well. The RBS masquerades as a legitimate base station to facilitate a Man-in-The-Middle (MiTM) attack between the mobile user equipment (UE) and the mobile network. An attacker can use the RBS to launch different attacks on mobile users and networks. These attacks include stealing user information, tampering with transmitted information, tracking users, compromising user privacy, or causing DoS for 5G services. The RBS threat has existed since GSM networks, and continued to evolve and persist with

the evolution of mobile networks. 5G networks are expected to introduce several security enhancements over 4G and legacy networks. Despite these security enhancements, 5G networks could still be a target to RBS-based threats using, for example, the following threat vectors:

- In a 5G Non-Standalone network, an attacker can exploit 4G interworking requirement to launch a downgrade attack to 3G and/or 2G;
- In a 5G Standalone network, an attacker can launch a downgrade attack to 4G then leverage the downgrade vulnerability in 4G to downgrade to 3G/2G;
- A compromised 5G small cell can create an RBS threat to 5G networks and customers causing a denial-of-service attack and/or perpetrate a MiTM attack;
- An attacker can exploit a lack of gNB authentication in an idle mode to force the user to camp on an RBS, which could lead to a denial-of-services (such as public safety warnings, incoming emergency calls, real-time application server push services, and etc.)

The following 5G vulnerabilities will be described in the following:

- Radio jamming is the deliberate jamming, blocking or interference with authorized wireless networks. A radio jammer is a transmitter that tunes to the same frequency and modulation as the opponents' receiving equipment with enough power to override any signal at the receiver.
    - The RAN OEMs need to develop new intelligence that can proactively alert the operator when this attack is initiated so that the operator can take appropriate actions to mitigate.
- Radio Sniffing techniques help to decode all sorts of essential network configuration details easily, and with low-cost software radios. Sniffing information can aid attackers in optimizing and crafting attacks.
    - This is a passive attack and will be more difficult on 5G radios operating in mmWave bands because most commercially available SDRs do not support the mmWave bands.
    - There are some Control Plane and User Plane vulnerabilities that will be present in 5G that 3GPP needs to address:
        - There are some Non-Access Stratum (aka Control Plane) vulnerabilities that were published in 2018 (i.e. LTEInspector [18]). These need to be addressed within 3GPP and the OEMs to mitigate these vulnerabilities.
        - There are some known User Plane vulnerabilities that relate to User Plane Key Reuse. The 3GPP needs to implement a UP key reuse restriction and more frequent refresh.
- RF spoofing refers to transmitting a fake signal meant to present as an actual signal.
    - Essentially this is an RBS transmitting and receiving on the licensed frequency bands of a particular operator.
    - In this scenario, the UE needs to be able to validate the legitimacy of the base station as being one owned and operated by the operator.
    - 3GPP has proposed in a study to use Digital Signatures to mitigate this threat but there has been no agreement on this to date.
    - The RAN OEMs need to develop new intelligence that can proactively alert the operator when this attack is initiated so that the operator can take appropriate actions to mitigate.

3. Mitigation and Recommendation Strategies

3. Mitigation & Recommendation Strategies

3.1 Significance and Security Risks

Sections 2 and 3 of this paper have outlined many of the threats, vulnerabilities, and considerations for 5G networks. This section will focus on some of the mitigations and recommendations associated with what was discussed. It should be noted that this is not a comprehensive list of mitigations but is meant to build upon the extensive work in the community found is various other publications.

3.1.1 Software for Open-Source in 5G

The 5G operator and vendor communities gain several benefits from adopting open-source software and are taking several actions to mitigate threats from this adoption.

Making source code available to the public significantly aids defenders, not just attackers. Continuous and broad peer-review, enabled by publicly available source code, improves software reliability, and security through the identification and elimination of defects that might otherwise go unrecognized by the core development team. Use of commercially available software, proprietary or open-source software, creates the risk of executing malicious code embedded in the software. Open-source software projects have a "trusted repository" that only certain developers (the "trusted developers") can directly modify. Since the source code is publicly released, anyone can review it or introduce malicious code.

For larger enterprises with multiple and vast repositories of code, it can be difficult to identify all the applications where open-source vulnerabilities may exist. Addressing the identification and mitigation challenge requires an intentional effort that includes activities such as code inspection, dynamic security scanning and vulnerability testing. These are the same techniques that should be applied to all software code repositories, open-source or not. There are also enterprise specific products that offer a complete end-to-end solution for third party components and supply chain management with features such as licensing, security, inventory, and policy enforcement. These products are offered by vendors such as Black Duck Software, Sonatype Nexus, and Protecode, to name a few.

Searches of Mitre CVE and NIST National Vulnerability Database (NVD) for vulnerability information provide limited information on open-source software vulnerabilities. Information on open-source software vulnerabilities is distributed among many different sources and is hard to track. To address the risk of open-source vulnerabilities in the supply chain, groups such as PCI and Open Web Application Security Project (OWASP) have specific controls and policy in place to govern the use of open-source components. Other security repositories exist, including the Node Security Project for JavaScript/Node.js specific vulnerabilities and Rubysec for Ruby specific vulnerabilities. However, there are still many open-source projects and ecosystems that are not well covered.

An entire market of open source and commercial tools has emerged over the years to tackle this problem as a result. These tools vary in approach and capabilities, and some are open-source themselves. Most of these tools use the NIST NVD as a starting point for sourcing open-source software vulnerabilities. Each tool is then enhanced with usability features and/or additional data sourcing for improved functionality. In general, an open-source software security analysis should:

- check for public vulnerabilities (ensure the open-source components do not contain publicly known vulnerabilities, reported with vulnerabilities described in other public resources);
- use commercial security intelligence (use additional vulnerability data sources, such as from data vendors, to augment the public vulnerability data);
- perform static analysis (use static analysis tools to validate that the open-source components do not contain unreported security vulnerabilities);
- perform comprehensive security reviews (perform a comprehensive security review of the open-source component).

3.2 Zero-Trust Security

In this age of rapid adoption of connected devices and nearly ubiquitous connectivity, a Zero-Trust model will help address security concerns for 5G and beyond. Security in 5G introduces solid protection controls. For instance, 5G introduces an improved encryption process which improves anti-tracking and spoofing features, making it difficult to monitor and track connected devices. Another example is the introduction of software and cloud-focused solutions that quickly detect and mitigate threats. No system is perfect; this is highlighted by the fact that Stingray devices can still be used to deploy fake base station attacks which enable bad actors to intercept user traffic, and possibly tamper with the data. Zero-Trust security for 5G is a technique to further help mitigate security concerns.

5G brings about virtualization, and a network operator might opt to run slices of its virtual network functions on an external cloud infrastructure while other slices run on the network operator systems. In this case a new actor is the external cloud provider who is not part of the existing network operator trust model. 5G MEC deployment models will include private cloud, public cloud, and hybrid models. External cloud providers not part of the existing network operator trust model will be a new actor in the public and hybrid cloud models. The external cloud provider might also have data centers in various jurisdictions, and it is not always clear which jurisdiction the virtual network functions maybe running in. This concept can be referred to as "5G without borders," which could create trust concerns. In this scenario, Zero-Trust is even more important than it may have been in pre-5G, and security controls such as policy enforcement of geo-location and mutual authentication of entities where virtual network functions are running are obligatory to establish the needed verification.

Zero-trust security ensures that security is in place from untrusted domains (e.g., supply chain, Internet, user devices, other operators and partners) to and from within trusted domains (operator networks). Operators must acknowledge that Zero-Trust can be considered to mitigate security threats in 5G.

3.2.1 Physical Aspects

Zero-Trust cannot be achieved without the full participation of all the physical elements in the trust chain in the network. As 5G infrastructure grows, the attack surface continues to expand, making it difficult to define the boundaries of the surface. To highlight, like 4G, 5G is not going to be a flash cut. Instead, 5G will evolve side-by-side with 4G, with physical evolutionary phases taking place over the next decade. Thus, 4G and 5G physical elements will coexist for some time. With Zero-Trust, rather than focusing on the macro-level of the attack surface, the protection surface needs to be determined. Migration from physical elements in the network with relatively untrusted implementations will be replaced with physical elements with substantially more trust built into the product(s) from the supply chain. Some examples of building Supply Chain Trust are:

- **Trusted HW**: Silicon selection from audited sources and Certificate protection should be built into Device Product lifecycles with a hardware root of trust. Device Supply Chain Security/PKI should include anti-cloning, multi-layer security with end-to-end identity/PKI protection, traceability, and reporting.
- **Trusted Certificate Authorities for Identity Provisioning**: OEM/ODM Certificates should be audited by recognized auditing entities (such as WebTrust [19]). The CA should be required to support Certificate revocation by vendor and all devices should support renewing/updating existing certificates in the field.
- **Trusted SW**: Software should be secured with a robust, stack-based code signing approach. It needs to cover the bootloader, OS and applications at time of deployment, as well as authenticated version upgrades to make it more difficult to introduce malicious software into operator-controlled elements. The code signing system should provide for reporting and traceability to allow for tracing sources of malicious code introduction to an individual if introduced through supply chain or operator provided software. Code Signing Features for testing and debugging should severely limit global back-door opportunities such as unlocking devices individually for a limited period or number of reboots.

Operationally, the protection surface will have to encompass the critical physical systems that are most valuable to the network and the organization. The physical infrastructure is set to evolve and will no longer be within defined perimeters. Hence, more sophisticated tools will need to be considered to offset this perimeterless dilemma by ensuring that all physical systems are accounted for, including internal and external systems. An inventory management system must account for all the physical elements that the operator owns on premise and in the cloud. Additionally, in the cloud, advanced firewalls, Unified Threat Management (UTM), Virtual Private Networks (VPN) and captive portals may also be used to help harden the perimeter around the telco fortress.

3.2.2 Logical Aspects

Application or software trust falls within the logical aspect of Zero-Trust. Software aspects such as code and license attributes need to be continuously monitored and validated. For example, a new form of remote integrity monitoring of Linux Ring 0 is one option to improve the state of software trust in a complex software platform that spans UE, RAN, Core, Application components and beyond. In addition, the need for a logical Zero-Trust Security Policy is equally important. This security policy should encompass enforcement rules of the physical policy as additional security measures as well as a mapping between the physical and logical rules. The security policy serves as an extra layer of security to govern how virtual elements communicate with physical elements. Furthermore, mapping the logical to the physical footprint is needed to provide a clear picture of how applications are run, how the traffic moves across the network and how it can be further segmented at the logical layer. Segmentation will include, but is not limited to data, applications, assets and services. Each of these elements will have its corresponding policy statements that are limited to the specific functions they are responsible for providing.

3.2.3 Operational Aspects

The way traffic moves across the network determines how it should be protected. Thus, the operational imperative is to gain contextual insight around the interdependencies of the network. This will also include edge locations scattered across multiple geographical areas and how it is inter-connected with the rest of the network. Understanding how edge resources interact allows for proper enforcement of security controls and provides valuable context to ensure the controls help protect both the network and the edge.

To fully realize the Zero-Trust principle, operators must put in place the proper operations that can monitor both physical and virtual infrastructure on an ongoing basis. The fact that the network in 5G will expand to the edge and will eventually be accessed by a multitude of connected devices makes security even more complex, and will more likely overwhelm traditional security model. Hence, monitoring and maintaining the network including the edge is a must. This approach includes inspecting and reviewing all logs, internal and external, including all layers while focusing on the operational aspects of Zero-Trust. Since Zero-Trust is an iterative process, inspecting and logging all traffic will provide valuable insights into how to improve the network operations including both centralized and distributed locations.

3.3 Cyber Threat Intelligence for 5G

The 5G architecture introduces new threat vectors, but it also provides inherent advantages to build a more secure network and user experience. Threat Intelligence is vital to stopping Threat Actors by identifying and mitigating threats before the threat impacts network availability and service delivery. Threat Intelligence is a crucial aspect of a strong 5G security posture.

3.3.1 Cyber Threat Intelligence Overview

The key terms for Threat Intelligence are:

- **Cyber Threat Intelligence (CTI):** Intelligence gathering of cyberthreat information to identify and profile threat actors and cyberattacks.
- **Indicators of Compromise (IoC):** Information about known cyberthreats and threat actors, including compromised or suspicious IP addresses, domains, URLs, as well as signatures for known malware, traffic patterns and behavioral or reputation characteristics.
- **Automated Threat Intelligence (ATI):** Automated methods to detect and mitigate cyberthreats through traffic analysis without the need to rely upon external data feeds.
- **Zero-Day:** A cyberattack utilizing an unknown vulnerability for which there is no available information, patch, or signature.

Threat intelligence is evolving from Static Configuration to Reactive protection to Predictive protection. Indicators of Compromise (IoC) provide reputation-based protection to known threats, but the dependence upon databases and static information exposes risk to zero-day attacks. Automated Threat Intelligence (ATI) using heuristics, behavior analysis, anomaly detection, and supervised/unsupervised machine learning provide zero-day protection, faster detection, and fewer false positives without the need to update intelligence feeds. ATI and IoC can be combined to provide multiple layers of threat analysis, as shown in Figure 10, for fast detection and mitigation of zero-days and known threats. A brief description of each threat intelligence layer is also listed.

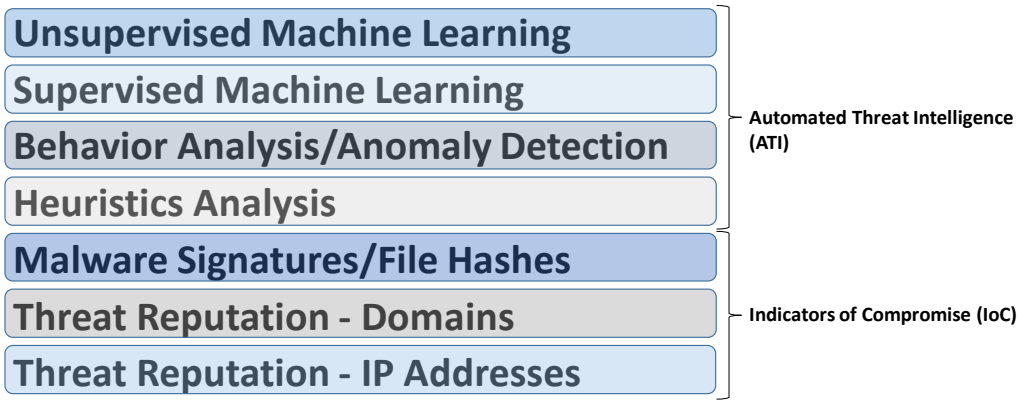


Figure 10: Threat Intelligence Layers

- **Unsupervised Machine Learning:** Uses pattern analysis and behavior analysis without inputs.
- **Supervised Machine Learning:** Uses pattern analysis and behavior analysis with input of data sets.
- **Behavior Analysis/Anomaly Detection:** Establishes a baseline for known and expected behavior for devices and IP addresses. Apply this learned knowledge to detect abnormal behavior due to malfunction, compromise, or infection.
- **Heuristics Analysis:** Analyze payloads for known suspicious content.
- **Malware Signatures/File Hashes:** The database of known malicious code "in the wild" that is applied to payloads to detect the malicious code as it propagates.
- **Threat Reputation Lists:** List of known malicious websites, hosts, and command and control servers identified by domains, URLs, or IP addresses.

3.3.2 Cyber Threat Intelligence Use Cases for 5G Security

Threat Intelligence can be applied to a wide variety of 5G attack vectors to reduce and limit the attack surface on the User Plane and Control Plane.

- **User Plane**
  - Identify device-type (including IoT device type) and behavior profile to detect and mitigate an internal botnet targeting network function(s) with the purpose to cause Denial-of-Service (DoS) in the 5GC (for 5G SA) or 4G EPC (for 5G NSA). The 4G EPC, which uses Diameter signaling and is built for 4G density, is particularly at risk of a massive IoT botnet from the 5G NR when operating 5G NSA.
  - Identify device-type (including IoT device type) and behavior profile to detect and mitigate internal botnet targeting public internet with purpose to compromise operator's reputation.
  - Detect and mitigate inbound malware sourced from external systems targeting network functions. 5G Fixed Wireless Access (FWA) routers may be at risk due to recent attacks targeting home routers conducted by nation-state threat actors. Examples are the VPNFilter attack and the COVID-19 Oski attack. The severity of this risk has gained the attention of Senator Mark Warner-VA, who serves on the Senate Intelligence Committee [20].
  - Detect and mitigate inbound malware source from external systems targeting User Equipment (UE). IoT devices are at particular risk of malware infection due to limited built-in security capability, the wide variety of known existing threats, and the constant introduction of zero-day attacks.
  - Detect and mitigate external volumetric DDoS attack targeting network resources with purpose to cause outage or degrade end-user quality of experience.

- **Control Plane**
  - Detect and mitigate signaling storm to protect 5GC functions due to internal massive IoT botnet so that the network remains available and there is no interruption to services.
  - Detect and mitigate SS7/Diameter signaling storm to protect 5GC functions from DdoS when roaming on 4G RAN.
  - Detect and mitigate DNS flood attack while passing valid DNS messages. The ability to identify the attack source and profile the signature of the attack enables valid DNS messages to be permitted so that a network outage can be prevented.
  - Detect and mitigate BGP flood attack while passing valid BGP messages. The ability to identify the attack source and profile the signature of the attack enables valid BGP messages to be permitted so that a network outage can be prevented.

3.3.3 Cyber Threat Intelligence in the 5G Architecture

Network operators can reduce the attack surface using Threat Intelligence with four key concepts in the 5G architecture. These are Multi-access Edge Compute, Data Network (DN), Network Slicing, and Virtualization.

1. **MEC deployment**
   - Closer to the RAN and UEs to isolate and mitigate internal Botnets and DDoS attacks
   - Separation from Core to isolate attacks
   - Scales for UE Density

2. **DN deployment**
   - Closer to Public Internet to isolate and mitigate external volumetric DDoS attacks and block malicious payloads.
   - Separation of LAN services from Core

3. **Network Slicing**
   - Optimizes ATI. Heuristics, anomaly detection, and machine learning are applied to traffic sets, reducing false positives.

4. **Virtualization**
   - Service Chaining to security services can be applied per application and per enterprise
   - Scales for Massive IoT
   - Orchestration provides dynamic instantiation of VNFs to respond to security incidents. Security resources can be dynamically allocated for surges in bandwidth, signaling rates, and UE density due to cyberattacks. Orchestration of ATI delivers faster Incident Response (IR).

3.4 5G Security Capabilities

3.4.1 Security Architecture

This section posits an overview of some of the most important security capabilities offered by 3GPP standards to enable confidentiality, integrity, availability, authentication, privacy, isolation and resilience in 5G.

The 3GPP SA3 Group has defined the security architecture of 5G mobile networks in specification document TS 33.501. The goal is to improve the security state of 5G far beyond its predecessor generation networks. The security architecture is multidimensional, and operators must employ a continual risk-based approach to deploying, operating, and monitoring network and services. Operators must also evolve security controls around emerging security threats. Table 2 provides a glimpse of the security dimensions which need to be considered in the security architecture.

Table 2: 5G Security Dimensions [21]

Security Dimensions	Description
Network Access	UE authentication and access (3GPP, non-3GPP)
Network	Control plane security, user plane security
Device	Device security capabilities (USIM)
Application	Application workloads at edge and central locations in the application domain have to be protected (user and provider)
Service Based Architecture (SBA)	Authentication and transport security protection between network functions, authorization framework, and improved interconnect security.
Visibility & Monitoring	Full visibility can help achieve end-to-end security

3.4.2 Security Functions

3.4.2.1 Enhanced Confidentiality

Confidentiality protection of user data, RRC signaling, and NAS signaling are optional. Operators are urged to implement and enable confidentiality protection capabilities when applicable. In 5G, the base-station is logically split in the interface between the CU and DU elements. Security is provided for the CU-DU interface. The DUs are typically deployed at some edge of the network, and do not have access to any user data when confidentiality protection is enabled.

3.4.2.2 Enhanced Integrity

Prior to 5G, mobile architecture did not have a trust model using integrity protection at the User Plane. 5G enables the ability to introduce integrity protection of the User Plane which renders any mutual authentication susceptible under the stress of attacks. One of the key security considerations in 5G pushes for a change in the overall architecture and the trust model leveraging the security principle of key segregation, and that User Plane integrity protection will be supported. In other words, integrity protection for user data must be enabled at the UE and gNB, but their use is optional. Integrity protection of user data adds overhead on packet sizes and increases processing load at both the gNB and the UE. Nevertheless, enabling integrity protection provides augmented security protection for user data. Furthermore, for NAS and Radio Resource Control (RRC) signaling, integrity protection is mandatory.

At the time of writing this white paper, there is no consensus to make integrity protection mandatory in 5G standards for full rate services. This will result in some UEs and networks being deployed without this capability (or limited to very low rate services), and it will be difficult to close off this gap in the future. However, operators may consider enabling the integrity feature gradually to accommodate for key metrics such as performance, latency, and bandwidth. The gradual enablement of the integrity feature can be parametrized

to include different conditions such as limiting its feature use for specific device types and, amount of bandwidth to prevent any network degradation issues.

3.4.2.3 Enhanced Authentication

Authentication is of critical importance to cellular networks because they form the foundation for protecting users, networks, services, and communication between them. 5G authentication enhances 4G authentication from different perspectives, including a unified authentication framework, better UE identity protection, enhanced home-network control, and more key separation in key derivation. In 5G, the primary protection of UE traffic is assured between the UE and the AMF. A secondary protection is based on Extensible Authentication Protocol [22] (EAP) which is established between the UE and the AAA server in external data networks, where the core network element SMF plays the role of an authenticator. This secondary protection is optional but encouraged for the user plane.

AMF triggers primary authentication of the UE using the Subscription Concealed Identifier, which is a one -time temporary user identifier. Then, it assigns a 5G-GUTI to the UE and supports reallocating the 5G-GUTI to the UE when necessary (e.g., roaming). This is performed on both 3GPP (e.g., gNB) and non-3GPP (e.g., WiFi) access. The Security Anchor Function (SEAF) at the AMF performs primary authentication procedures and stores the security profile on a per subscriber basis for the duration of the registration. The IMSI has not managed to keep its name in 5G, as is the case with previous generations of mobile networks. Rather than the IMSI in 5G, the SUPI is used for UE authentication and key agreement. The SUPI is transferred in a cipher text form over the 5G RAN.

Security Edge Protection Proxy performs mutual authentication and negotiation of cipher suites in the roaming network. This prevents third party operators' devices from tampering with sensitive data such as user IDs exchanged between core networks.

3.4.2.4 Threat Mitigation for 5G SMS over NAS Interoperability

As described in Section 2.3.8, 5G SMS over NAS may expose 5G network operators to SS7 or Diameter vulnerabilities when a 5G network operator interoperates with their legacy 3G/4G SMS systems or interconnects through an inter-carrier IPX provider.

The exposed 5G network operator can benefit by following GSMA recommendations on remediation of SS7 and Diameter vulnerabilities. Central to this remediation are the GSMA recommendations to implement an SS7 or Diameter firewall at the MAP interface on the UDM, and to filter/block malicious SS7 or Diameter traffic (described in GSMA FS.11 SS7 Interconnect Security Monitoring and Firewall Guidelines and GSMA FS.19 Diameter Interconnect Security). These documents offer further mitigation guidance on risk assessments, firewall rules, traffic monitoring, potential abnormalities to look for, and the means to report them.

3.4.2.5 Enhanced Privacy

In 5G, subscriber privacy is improved by ensuring that the subscriber identifier (SUPI) is encrypted (as described in Section 2.3.6), and mutual authentication exists between the UE and the network element. The SUPI is used only inside the network, and may either be based on an IMSI, as defined in 3GPP TS 23.003, or be used as a network specific identifier user for a private network. As previously mentioned, there are a few cases where SUPI is sent in the clear. Operators should strive to reduce scenarios like customers connecting BYOD devices to the network before they have been updated with 5G credentials.

The EU General Data Protection Regulation (GDPR) may have legal implications which require both vendors (when building devices and products) and operators (when deploying and operating networks and offering services) address subscribers' rights such as consent and data portability.

3.4.2.6 Anti-Bidding-down Between Architectures

5G has also defined the ABBA (Anti-Bidding-down Between Architectures) parameter. This parameter allows the 5G system to enforce that a UE cannot access the network using older mechanisms that have had vulnerabilities associated with them. The value is currently defined from 3GPP release 15 as zero essentially represents a security version. As the 5G system adds new security algorithms over time, the UEs can know the full suite of security capabilities that are needed to attach to the network.

However, it could be until 3GPP Release 17 or 18 before the value needs to be incremented. ABBA is designed to be used in the future. Currently there has been no reason identified in release 16 to increment the value. The UEs should check to see that the value is zero in order to attach to the 5G system.

3.5 Slicing

As stated in Section 2.1.2, Network Slicing is the technique of isolating the end-to-end performance of a portion of the network compared to another. Slicing enables a large set of solutions in a scalable manner.

Security isolation is achieved through the proper isolation of both physical and logical networks. Operators can configure the physical network with multiple logical networks. Operators can also configure multiple logical networks with various network features (e.g., namespaces, VLANs). In the 5G context, each network slice can serve an application or a set of applications which can provide one or a mixed set of services (e.g., voice communication, video streaming or Internet of Things). In addition to physical and logical segregation, traffic and resource isolation are also critical. Traffic isolation can be provided by creating specific virtual lanes to serve specific traffic types that can be routed through a specific logical/physical lane pair. Resource isolation should be supported in the underlying infrastructure. Isolating traffic and resources play an important role in safeguarding critical information systems from security vulnerabilities and attacks. For instance, the scope of a potential attack that can be restricted to a particular slice should such lane be compromised. Although virtual slices can serve as a good example of isolation, it is prudent to note that isolation configuration in terms of the number of slices, the amount of traffic and resources that go into the slice and the slice security setup must be tailored to the specific use case(s) being served.

3.5.1 What Slicing Can Enable and Provide

When slicing is used it can support the following enabling functions:

- **Physical Resource Separation**
  - Over the air via 5GQoS (Different channels or at resource block level)
  - Over the air with dedicated spectrum (for instance dedicated public safety)
  - Core Network Element functions (dedicated network elements can be reserved for particular slices)
- **Traffic Separation:** both in the RAN and with mappings transit networks (e.g. VLANs, MPLS, Tunnels).
- **Custom Routing/Steering of traffic:** to support needs of end customers. Virtualization of 5G Core comes with Virtualized networks, which allows easy integration to support other cloud environments, Steering and VPN functions for enterprise customers. This steering can also be leveraged for service chaining.
- **Scalability:** many of these functions were done in the past based upon APNs (now termed DNNs in 5G). Adding slices greatly improves the scalability. In addition, private networks work in 3GPP has created the Network Identifier (NID), something similar to an APN/DNN but is based on the NAI type which has unlimited scale. Slices also inherit this scale as they can be defined/managed per-NID and per-DNN.
- **Optimization:** the 5G standards also offer more formality and standardization of optimizations in the network (e.g. caching of per UE per NID/DNN slices in AMF).

These enabling functions combined with other 5G technologies are used to provide MEC, URLLC, V2X, time sensitive networking (TSN) and more. Many of these are needed for Industry 4.0, the combination of IoT and automated manufacturing envisioned for a lights-out factory. Scalability and optimization have been missing from previous solutions which were created piece-meal over 15 years in an ad-hoc fashion.

3.5.2 Basics of How Slicing Works

Figure 11 details the non-roaming reference architecture from 3GPP 23.501. Marked in blue are some key interface points used when a UE registers and wants to use slices. Figure 12 (SECaaS built upon Network Slicing) is also a good visualization. The descriptions below give a basic explanation of how these interfaces are used to assign network resource functions to one slice the UE wants to use.

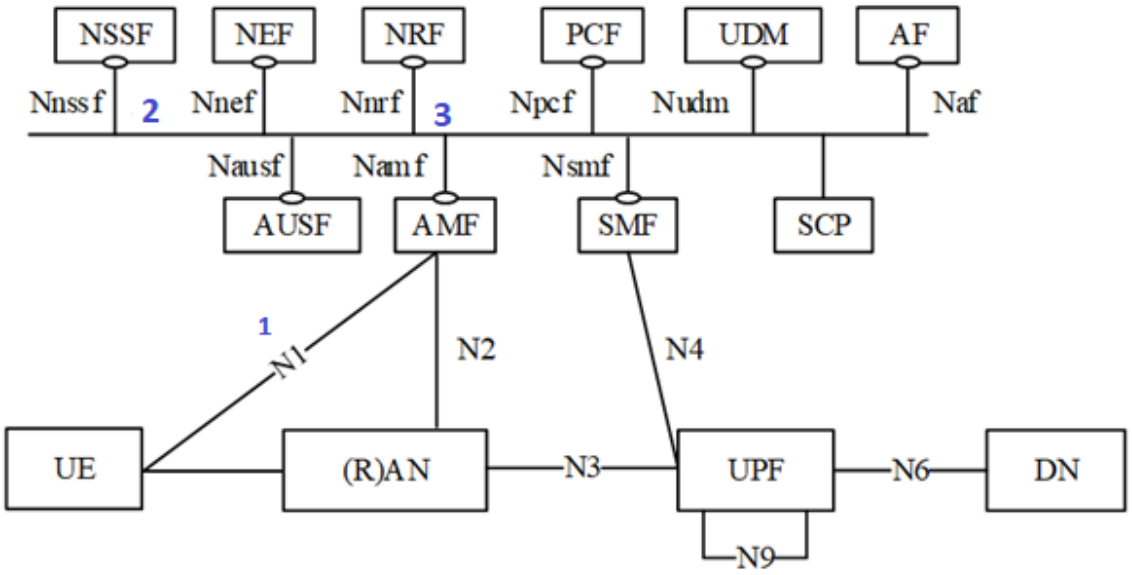


Figure 11: Non-roaming reference architecture from 3GPP 23.501

1. The UE/mobile is essentially either pre-provisioned or told by the network which slices and DNNs it is to use. When the UE first contacts the network it is assigned an initial AMF by the 5G RAN and it indicates to the default AMF it wants to use these slices. The RAN is configured with a default AMF corresponding to each PLMN and NID. The UEs possess the pubic root certificate and full chain of each PLMN and/or NID they can access. The UEs encrypt sensitive information with the pubic key of the PLMN or NID.
2. The initial AMF knows some slices the mobile wants to use but in this case, it does not know one of the slices that particular UE wants to use. The Initial AMF checks with an NSSF (Network Slice Selection Function) which knows a list of NRFs that will know an AMF that should handle the slice for the UE. The interface model for this information exchange is defined in 3GPP TS 29.531. The query parameters include, Slice, Tracking Area (Location) and SUPI.
3. After the AMF finds out the correct AMF to use from an NRF it redirects the mobile to use that AMF for that slice. The interface model to NRF is defined in 3GPP TS 29.510. The query parameters include, Slice, Tracking Area (Location) and SUPI.

The selected AMF could be very close to the UE. Here are some service examples indicating where that AMF may physically reside given a set of deployment use cases:

- AMF could be a V2X AMF close to the UE (UE could be a car that was just turned on)
- AMF could be an AMF in a Private Network or, in standard notation, a PNI-NPN (Public Network Integration – Non-Public Network)
  - Enterprise User on Private Network – Campus Site,
  - Manufacturing Center (untethered robot)
  - Mine (robot truck, raw material processing machine)
  - Farm (tractor)
  - Car Dealership (AR Goggles for technician in shop, a car to perform Firmware download)
  - Naval Port or Refinery (UE is an inspection drone)
  - Near a convention center for Public Safety (UE is headset for fire fighter)
  - Dedicated AMF cloud pool for power company (UE is power meter or inspection drone)

The SMF (Session Management Function), UPF (User Plane Function) and other network functions are also discovered and assigned in a similar manner based upon a slice. All of these elements can be located very close to the UE in order to facilitate low latency communications given network points of presence for the solution at hand.

In each case above, the slice is leveraged to separate and steer the user traffic to an Edge or private network for each solution. The slice can be thought of as an extension of that network. Slices can also be used to provide levels of service such as 5G QoS, 5GLAN Service, TSN Service and steer traffic onto RF bands.

Other important attributes of the 5G System related to slices include:

- **Optimizations:** The AMF can cache up to 8 slices for each UE. If the UE re-registers it does not need to go through the whole selection process again as the AMF used can handle the slice (i.e., it is optimized).
- **Additional Optimizations:** The AMF is notified if a UE can no longer use a slice automatically by the NSSF (no data dips into the core to check every time).
- **Enterprise Access Control:** Each slice can support a 2nd authentication and authorization in addition to the mobile provider authorization. (The enterprise has control over access to their network defined by the slice.)
- The example above is only for UE registration. Similar exchanges and resource function assignments are done during mobility events and other events.

3.5.3 Security for Slices

The integrity of the interfaces mentioned above are very important from a security perspective. Compromise of either of these interfaces could lead to user traffic being sent to the wrong network. Great care should be taken to ensure the fidelity of these interfaces and network functions using the interface. It is also paramount that the slices are assigned to the correct subscribers NIDS and DNNs, and fulfillment and Slice Management are crucial as well. Chapter 15 of 3GPP TS 33.501 is dedicated to Slice Management. It specifies the use of TLS and mutual authentication between entities.

The security sections in the 3GPP SBI interfaces models refer to 3GPP TS 33.501 for OAuth2 authorization and mandate 2-way authentication. 3GPP TS 33.501 section 13.1.0 details general requirements. TLS and certificates are mandated but flexibility left open for other secure options, for example, NDS (IPSEC). Some of the specs indicate that confidentiality is not mandatory if function's interfaces are on a physically secure network. Mutual authentication appears to be mandated on all interfaces. The NRF model defines a discovery service as well.

3.5.4 Privacy and Slices

The examples above are relatively simple. There are many types of slice identifiers each with sub-parts to

handle privacy and security. Some are used only in the home network core, roaming, or some are encrypted in the clear. Slice and Network identifier management is also customizable to a certain extent by each provider.

This Slice and NID management should be accomplished in a manner to meet the following subscriber privacy requirements:

1. NID, DNN and Slice identifiers should not be sent in the clear over the air to thwart identification of any organization affiliation with a subscriber.
2. Some enterprise customers may want their organization affiliation hidden from a set of visited PLMNs but still have the subscriber receive the services while roaming.

Recommendations:

- **DNS Security:** Use integrity protected DNS for network function communications. With the introduction of private networks, confidentiality of the DNS might also be warranted. DNSSEC and DNS over TLS/DTLS are options, but signed records may be all that is needed. Some NFs, such as NSF, have a discovery service associated with them but if DNS is used to find the discovery service IP the same threat exists.
- **SBI Interface Security (TLS):** Mandating mutual authentication, confidentiality, and integrity protection (e.g. TLS) over the Service Based Interfaces (SBIs)—regardless of physical network access—is a method to help protect against insider or physical access breaches. These interfaces typically contain subscriber information (SUPI (NID, DNN)), location (TA), and List of Slices (NSSAI—possibly organizational affiliation if slice mapping is available).
- **SBI Interface Security (NDS):** Using NDS (3GPP Network Domain Security) alone without TLS is likely only applicable if only a few NFs are remote. IPSEC would need to terminate on each remote host (client mode) to the SeGW to be secure. Using an SeGW-to-SeGW site-to-site solution would still have confidentiality issues locally at the remote site. Physical access control would be needed. Native IPSEC could be leveraged in IPv6 capable networks but most providers don't want to expose such functions' addresses directly to a non-provider network.
- **Private Network SBI Interface Policy Controls:** With the introduction of Private Networks (when NID/slicing is used and certain NFs reside on the enterprise premises), the Oauth2 form of authentication may need to be re-assessed at least in terms of how a private network element might interface with the provider's core for Oauth2. Policy restrictions on interface operations could be put into place on some of these interfaces to mitigate these concerns while leaving Oauth2 in place as is. Oauth2 was designed to be implemented across multiple enterprises, and some best practices might need to be developed at least internal to each provider.
- **Privacy Considerations:** S-NSSAI, NID and DNN should not be in the clear over the air.
- Other Items: The same concerns, mitigations and potential enhancements pointed out in the rest of this paper exist for slicing as well. When slicing involves NFs located in non-provider networks new threats can emerge and there is less control of security policies. An operator should be mindful of placing equipment necessary for regulatory compliance in these situations.

3.5.5 Private Networks Security

3GPP has done well defining the security and privacy for handling slices while still leaving flexibility to implement more or less stringent security when interfaces are physically protected in the provider's network. The use of slices is an optional feature of the 5G system and many of the solutions provided by slicing can be provided by other means. Private Network customers and Edge Deployment in 3rd party transport networks will likely force the more stringent baseline security recommendations as needed if it makes sense for a provider to address these markets. Below are recommendations:

- The secure options for SBI interfaces should be used.
- NDS client mode recommended for RAN sharing.
- NID, NSSAI should be encrypted over the air.
- PLMN List, NID, Slice IDs should be managed appropriately for certain customer in roaming situations.
- Local NPN NF components should also follow secure options for interfaces.
- CAG Portals and APIs should support 2-way authentication.
- EAP authentication hop interfaces should provide confidentiality, integrity, and two-way authentication protection.
- For privacy MOCN and MORAN sharing solutions should be configured whereby each core accessed (default AMF) is by PLMN or NID with two-way authentication between RAN and Default AMF.
- GWCS solutions should have a trusted provider manage the default AMF. Not following this policy could violate privacy of public users in PNI-NPN deployment locations.
- UEs should be configured to not pass sensitive parameters in the clear. If the UE does not have the public key of the PLMN or NID, during an emergency call, this could be an exception.
- UE should authenticate the core PLMN or NID and reject core if authentication fails.
- Some providers may need sensitive parameters in the clear as part of initial activation for PKI management. This should be a one-time process and restricted after this first occurrence.

Please refer to Sections 3.6 and 4.5 for additional information.

3.6 Edge Data Security

More information is available in Sections 3.5, 3.6 and 4.5. There are many EDGE Data Security considerations including:

- **Sensitive Data (Subscriber Credentials/Legal Intercept Information):** Cached sensitive data must be encrypted and integrity protected with high strength.
- **Boilerplate and Incremental Audits:** Boilerplate Audit (Set of NFs + Initial Applications) for Edge sites should also performed in addition to core audits. As each application or NF is added or upgraded an additional audit should be done incrementally on affected components.
- **Signed Code:** Integrity of all code running on NFs and Applications should be required as well as attestation checks ongoing (secure boot/chain of trust).
- **Interface Security:** Mutual Authentication, encryption and integrity protection between all NFs both at edge and between core sites and RAN.
- **DNS Integrity:** DNS should have integrity protection for resolvers to leverage.

3.7 Design and Implementation

There are many new design and implantation functions being worked on by the industry. The following describer many of the new developments.

3.7.1 Security-as-a-Service for 5G

The 5GC SBA enables new service types and supports a wide variety of diversified service types associated with different technical requirements. 5G enables a variety of possible use cases, each with unique security requirements. 5G Service Exposure and Network Slicing architectures are tools for CSPs to offer differentiated services to their consumer and business customers. Service exposure enables marketplace offerings for security services to be created, and network slicing segments traffic to apply security services that meet the use case. The operator's Security-as-a-Service (SECaaS) offering is built upon Network Slicing, which provides the "Security" for the use case, and Service Exposure, which provides the "as-a-Service". Multi-Tenancy enables customers to use templates and instance models to build security functions into their applications.

Network slices are used to deploy services at the multi-access edge across a distributed cloud infrastructure. Network slices can be configured based upon the service-type (eMBB, mMTC, URLLC), customer, and application to provide the required latency, bandwidth, QoS, and security. While slices provide inherent security through segmentation, slices can also be used to provide additional security protection and security policies specific to the use case and customer requirements. Customers select the security services from the operator's library. Examples include:

- An MVNO can select Reputation-based Threat Protection and Anti-Botnet Protection for data-only devices.
- An enterprise can select Firewall and DDoS Protection service.
- An industrial manufacturer can select IoT Behavior Anomaly Detection
- A public safety organization could choose a combination of all security functions.

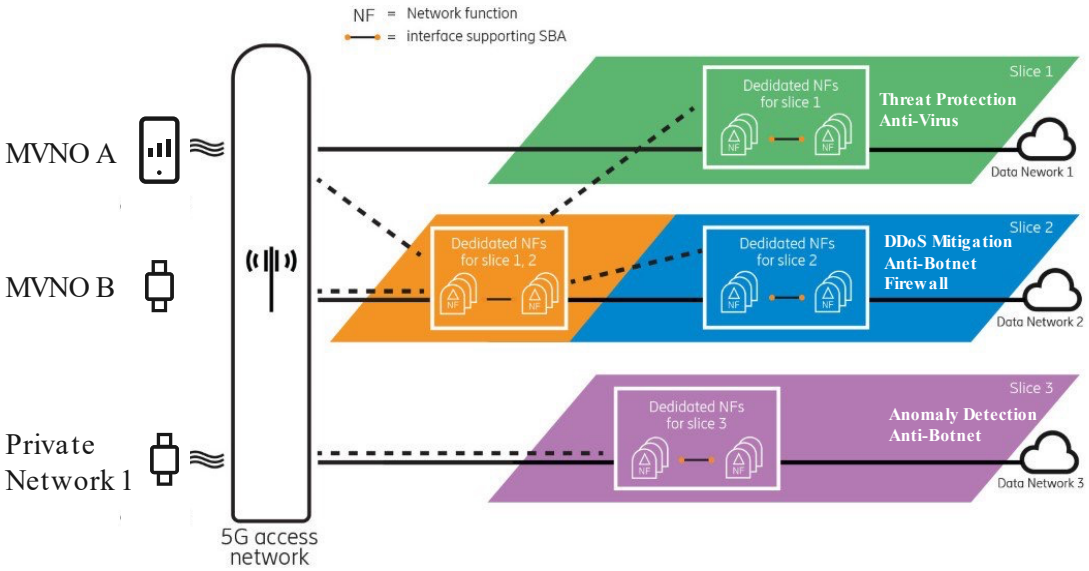


Figure 12: SECaaS built upon Network Slicing [23]

Service exposure and APIs enable application developers to activate libraries, functions, and containers through a cloud marketplace to discover and create solutions for internal use, consumers, business customers, and partners through an API Gateway. Service exposure makes it possible to interact with data streams to deliver microservices-based security functions. Service exposure covers [24]:

- **Standardized slice characteristics for pre-integration:** latency, bandwidth, proximity, QoS, rating
- **Value added services for differentiation:** security, geofencing, etc.
- **Instrumentation:** insights and monitoring

Service exposure will be critical to meet the security requirements of applications that rely on edge computing, network slicing and distributed cloud. Customers will be able to leverage the marketplace to create solutions with the appropriate security functions. The benefit to 5G customers is that applications are designed with the security functions to meet the use case. The benefit for the network operator is that SECaaS can be available directly from the network as a bundled security service for network differentiation, or as an opt-in security service for revenue-generation.

3.7.2 Zero-Trust

5G provides built-in security features, but they should not be viewed as the only options. The Zero-Trust model is essential to mitigate security risks. For instance, in 5G, each element should utilize a robust code signing stack at both the silicon and software layers. At the silicon layer, secure implementation of code signing should be in place. At the software layer, each element's code must be verified before it can be successfully loaded into the software stack. For instance, the firmware, operating system/hypervisor, and network function layers must be validated sequentially before the software stack layer can be fully enabled to serve network traffic and host data. Essentially, it is a stack of signed elements so that all the layers involved can be trusted, as illustrated in Figure 13.

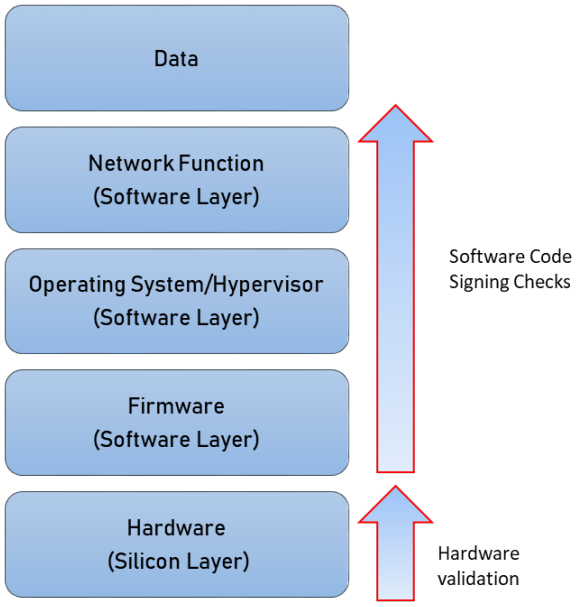


Figure 13: Zero-Trust Validation checks required at both the hardware and software layers.

Furthermore, Zero-Trust cannot be achieved without the full participation of all the elements in the trust chain for a network. Below are examples of common, but significant options to consider when adhering to Zero-Trust in digital systems:

- Use a True Random Number Generator (TRNG) in each appliance
- Use a trusted clock (or secure counter)
- Use secure storage
- Use of tamper-resistant devices from which the key cannot be extracted if compromised

To achieve Zero-Trust, operators need to adapt a default "deny all" mentality and start opening up network lanes and endpoints according to the least privileges' principle. In some 5G use cases, operators must get security right as traffic volumes and connected devices will be overwhelming. Born from a Zero-Trust mindset, multi-layered security significantly reduces risk posed by other parties and components in the messaging delivery chain, including those posed by hostile networks—5G or otherwise.

3.8 Key Take-aways

In this paper we've thoroughly examined what differentiates 5G from previous generations, reviewed threats from previous versions of mobile architectures upon which 5G is built, and looked at new threats and vulnerabilities in 5G.

The key take-aways are:

- 5G is a fundamentally different architecture than previous versions of mobile architectures. It is designed around a service-based architecture that supports cloud-native functions, disaggregation, open-source software and embraces automation, SDN and NFV. These new architectural elements can aid defenders, and not just become greater threats and vulnerabilities for attackers to exploit.
- Threats remain in the RAN but improved anti-tracking and spoofing features mitigate these threats.
- 5G network operators should address vulnerabilities from interworking with 3G/4G systems, including legacy vulnerabilities to LTE EPC for 5G NSA networks and for 5G SMS over NAS interoperability with 3G/4G systems.
- 3GPP has designed into 5G several improved security capabilities, including Confidentiality, Integrity, Authentication, 5G SMS, Privacy and Isolation.
- Network Slicing provides greater isolation, privacy, and greater security across 5G networks.
- Zero-Trust is not a "rip and replace" model, instead, it can augment existing architectures and adds to a defense-in-depth security strategy. Hence, it can be deployed iteratively while leveraging existing tools and technologies. Zero-Trust should be a fundamental security position for all 5G architectures and requires a multi-faceted approach employed to cover the physical, logical and operational aspects of the network.
- Cyber threat Intelligence gives 5G operators the ability to identify and stop bad actors before they cause harm.
- Through the service based architecture, 5G enables operators to offer Security-as-a-Service.

Conclusion

Conclusion

5G is a significant step forward for communications networks. Its ability to support massive bandwidth, massive interconnectivity of machines and reliable, low-latency communication will enable a variety of innovative applications, as well as one not yet envisaged. The new architectures that allow 5G to progress are also ones that can expose new vulnerabilities. Securing 5G must be designed-in and not be an afterthought. Hence, a careful approach to these new aspects of cloud-native services, open-source software, APIs, SDN and NFV can improve their security. Additionally, taking a Zero-Trust approach, combined with the advanced techniques of cyber threat intelligence, and Network Slicing that 5G offers will further enhance 5G's security. The transformation to a secure 5G will occur, however, its level of success will depend on making it deployable and operational.

Appendix

Appendix

Acronyms

3GPP	3rd Generation Partnership Project
4G	Fourth Generation
5G	Fifth Generation Mobile Networks
5GC	Fifth Generation Core
AI	Artificial Intelligence
AMF	Access and Mobility Management Function
BF	Beamforming
CA	Carrier Aggregation
CIoT	Cellular IoT
CLI	Cross-Link Interference
cMTC	Critical Machine Type Communications
CN	Core Network
CoMP	Coordinated Multi-Point Transmission and Reception
CP	Control Plane / Cyclic Prefix
CU	Control/ User Plane OR Central Unit
D2D	Device-to-Device
DC	Dual Connectivity
DCI	Downlink Control Indicator
DL	Downlink
DU	Distributed Unit
E2E	End-to-End
EB	Enhanced Beam forming
eMBB	Enhanced Mobile Broadband
EN-DC	E-UTRAN New Radio Dual Connectivity
eNodeB	Evolved NodeB
EPC	Evolved Packet Core also known as System Architecture Evolution (SAE)
EPC/SAE	Evolved Packet Core/System Architecture Evolutions
ePDG	Evolved Packet Data Gateway
EPS	Evolved Packet System
ETSI	European Telecommunications Standards Institute
FCC	Federal Communications Commission
FD	Frequency Division
FD	Full Dimension as in FD-MIMO
FDD	Frequency Division Duplex
FDM	Frequency-Division Multiplexing
FR1	Frequency Range 1 (410 MHz – 7125 MHz)
FR2	Frequency Range 2 (24250 MHz – 52600 MHz)
FWA	Fixed Wireless Access
GERAN	GSM EDGE Radio Access Network
GPRS	General Packet Radio Service
HSS	Home Subscriber Server
IAB	Integrated Access Backhaul
ID	Identity
IEEE	Institute of Electrical and Electronics Engineers
IMS	Internet Protocol Multimedia Subsystem
IMT	International Mobile Telecommunications
IIoT	Industrial Internet of Things
IoT	Internet of Things
IP	Internet Protocol
I-SMF	Intermediate SMF
ITU	International Telecommunications Union
LTE	Long Term Evolution
MEC	Multi-access Edge Computing
MIMO	Multiple-Input Multiple-Output
mMTC	Massive Machine Type Communications
mmWave	Millimeter Wave
MTC	Machine Type Communications
MU-MIMO	Multi-User Multiple-Input Multiple-Output
NEF	Network Exposure Function
NF	Network Functionality
NG	Next Generation
NID	Network ID
NPN	Non-Public Network
NR	New Radio
NRF	Network Repository Function
NR-U	NR Unlicensed
NSA	Non-Standalone
NSSAA	Network Slice-Specific Authentication and Authorization
OCB	Occupied Channel Bandwidth
OFDM	Orthogonal Frequency Division Multiplexing
OTA	Over-The-Air
P-CSCF	Proxy Call Session Control Function
PDCCH	Physical Downlink Control Channel
PDCP	Packet Data Convergence Protocol
PDSCH	Physical Downlink Shared Channel
PLMN	Public Land Mobile Network
PMCH	Physical Multicast Channel
PNI NPN	Public Network Indicated NPN
PRACH	Physical Random-Access Channel
ProSe	Proximity Services

PS	Packet Switched
PSBCH	Physical Sidelink Broadcast Channel
PSCCH	Physical Sidelink Control Channel
PSFCH	Physical Feedback Control Channel
PSSCH	Physical Sidelink Shared Channel
PTRS	Phase-Tracking Reference Signal
PUCCH	Physical Uplink Control Channel
PUSCH	Physical Uplink Shared Channel
QoS	Quality-of-Service
RACH	Random Access Channel
RACS	Radio Capabilities Signaling Optimization
RAN	Radio Access Network
RIM	Remote Interference Management
RIT	Radio Interface Technology (IMT-2020 proposal)
RMSI	Remaining Minimum System Information
RNTI	Radio Network Temporary Identity
RRC	Radio Resource Control
RRM	Radio Resource Management
RSRP	Reference Signal Received Power
RSRQ	Reference Signal Received Quality
RSSI	Received Signal Strength Indication
RSTD	Received Signal Time Differenfce
RTOA	Relative Time of Arrival
RTT	Round Trip Time
SA	Stand-Alone
SBA	Service-Based Architecture
Scell	Secondary cell
SCI	Sidelink Control Indicator
SCG	SeNB Cell Group
SCG	Secondary Cell Group
SCP	Service Communication Proxy
SC-PTM	Single-Cell Point-to-Multipoint
SFN	Single Frequency Network
SIM	Subscriber Identity Module
SINR	Signal-to-Interface-and-Noise Ratio
SL	Sidelink
SMF	Session Management Control Function
SNPN	Stand-Alone NPN
SON	Self-Optimizing or Self-Organizing Network
SRIT	Set of Radio Interface Technologie(s)
SRS	Sounding Reference Signal
SRVCC	Single Radio Voice Call Continuity
TA	Time Alignment
TDD	Time-Division Duplex
TDM	Time-Division Multiplexing

TDOA	Time Difference Of Arrival
TSC	Time Sensitive Communication
TTI	Transmit Time Travel
Tx	Transmit
UAV	Unmanned Ariel Vehicles
UC	UniCast
UCI	Uplink Control Indicator
UCMF	UE (radio) Capability Management Function
UDM	Unified Data Management
UE	User Equipment
UL	Uplink
UPF	User-Plane Function
URLLC	Ultra-Reliable Low Latency Communications
V2P	Vehicular-to-Pedestrian
V2V	Vehicular-to-Vehicular
V2X	Vehicle-to-Everything
VN	Virtual Network
WG	(3GPP) Working Group
WI	Work Item
WID	Work Item Description

Appendix

References

[1] 5G Americas, "The Evolution of Security in 5G," October 2018. [Online]. Available: https://www.5gamericas.org/wp-content/uploads/2019/07/5G_Americas_5G_Security_White_Paper_Final.pdf.

[2] 5G Americas, "The Evolution of Security in 5G: A 'Slice' of Mobile Threats," July 2019. [Online]. Available: https://www.5gamericas.org/wp-content/uploads/2019/08/5G-Security-White-Paper_8.15.pdf.

[3] "LTE (Communications)," Wikipedia, [Online]. Available: https://en.wikipedia.org/wiki/LTE_(telecommunication).

[4] "Setting the Scene for 5G: Opportunities & Challenges," ITU, [Online]. Available: https://www.itu.int/en/ITU-D/Documents/ITU_5G_REPORT-2018.pdf.

[5] "Should I Buy from This Site? How to Know if a Website is Secure," DigiCert, [Online]. Available: https://www.digicert.com/blog/buy-site-know-website-secure/.

[6] n.-V. v. s. r. a. c. s. v. c. i. l. s. n. s. a. S. w. b. w. i. t. o. h. s. c. a. m. Legacy (non-VoLTE).

[7] "Signalling System No. 7," Wikipedia, [Online]. Available: https://en.wikipedia.org/wiki/Signalling_System_No._7.

[8] "Diameter Protocol," Wikipedia, [Online]. Available: https://en.wikipedia.org/wiki/Diameter_(protocol).

[9] "ENISA - Signalling Security in Telecom SS7/Diameter/5G," Wikipedia, [Online]. Available: https://www.enisa.europa.eu/publications/signalling-security-in-telecom-ss7-diameter-5g/at_download/fullReport.

[10] "Old vulnerabilities could majorly impact 5G security," Wikipedia, [Online]. Available: https://www.itproportal.com/news/old-vulnerabilities-could-majorly-impact-5g-security/.

[11] " GSMA Network Equipment Security Assurance Scheme (NESAS)," GSMA, [Online]. Available: https://www.gsma.com/security/network-equipment-security-assurance-scheme/.

[12] "SOC for Supply Chain," AICPA, [Online]. Available: https://www.aicpa.org/interestareas/frc/assuranceadvisoryservices/soc-for-supply-chain.html.

[13] "Network Functions Virtualisation (NFV," ETSI, [Online]. Available: https://www.etsi.org/technologies/nfv.

[14] Open Networking Foundation, "SDN Architecture," Open Networking Foundation, [Online]. Available: https://www.opennetworking.org/wp-content/uploads/2013/02/TR_SDN_ARCH_1.0_06062014.pdf.

[15] "5G Security – Package 3: Mobile Edge Computing / Low Latency / Consistent User Experience," NGMN, February 2018. [Online]. Available: https://www.ngmn.org/publications/5g-security-package-3-mobile-edge-computing-low-latency-consistent-user-experience.html.

[16] 5G Americas, "5G at the Edge," [Online]. Available: https://www.5gamericas.org/5g-at-the-edge/.

[17] Diagram source: 5G Americas Member Contribution, [Online].

[18] "LTEInspector: A Systematic Approach for Adversarial Testing of 4G LTE, Purdue University and The University of Iowa.," [Online].

[19] "WebTrust seal program," CPA Canada, [Online]. Available: https://www.cpacanada.ca/en/business-and-accounting-resources/audit-and-assurance/overview-of-webtrust-services.

[20] "Warner Urges Internet Networking Device Vendors to Ensure Security of Consumer Internet Connectivity Products," Mark R. Warner US Senator from the Commonwealth of Virginia, [Online]. Available: https://www.warner.senate.gov/public/index.cfm/2020/3/warner-urges-internet-networking-device-vendors-to-ensure-security-of-consumer-internet-connectivity-products.

[21] 3GPP, "TS 33.501, Security architecture and procedures for 5G System".

[22] "IETF RFC 3748: "Extensible Authentication Protocol (EAP)".".

[23] Ericsson, "Adapted from Ericsson white paper "5G security – enabling a trustworthy 5G system"," [Online]. Available: https://www.ericsson.com/49s169/assets/local/reports-papers/white-papers/s30209935-04_wp_5g-security_editb-mar18.pdf.

[24] Ericsson, "Edge Computing and Deployment Strategies for Communication Service Providers," February 2020. [Online]. Available: https://www.ericsson.com/en/reports-and-papers/white-papers/edge-computing-and-deployment-strategies-for-communication-service-providers.

[25] "Setting the Scene for 5G: Opportunities & Challenges," ITU, [Online]. Available: https://www.itu.int/en/ITU-D/Documents/ITU_5G_REPORT-2018.pdf.

Acknowledgments

5G Americas' Mission Statement: 5G Americas facilitates and advocates for the advancement and transformation of LTE, 5G and beyond throughout the Americas.

5G Americas' Board of Governors members include AT&T, Cable & Wireless, Ciena, Cisco, CommScope, Crown Castle, Ericsson, Intel, Mavenir, Nokia, Qualcomm Incorporated, Samsung, Shaw Communications Inc., T-Mobile USA, Inc., Telefónica and WOM.

5G Americas would like to recognize the significant project leadership and important contributions of group leader David Krauss of Ciena, along with many representatives from member companies on 5G Americas' Board of Governors who participated in the development of this white paper.