SECURITY GUIDANCE FOR 5G CLOUD INFRASTRUCTURES

Part III:
Data Protection

2021

TLP:WHITE

DISCLAIMER OF ENDORSEMENT

The guidance in this document is provided "as is." In no event shall the United States Government be liable for any damages arising in any way out of the use of or reliance on this guidance. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not constitute or imply its endorsement, recommendation, or favoring by the United States Government, and this guidance shall not be used for advertising or product endorsement purposes. All trademarks are the property of their respective owners.

PURPOSE

NSA and CISA developed this document in furtherance of their respective cybersecurity missions, including their responsibilities to develop and issue cybersecurity specifications and mitigations. This information may be shared broadly to reach all appropriate stakeholders.

CONTACT

Client Requirements / Inquiries: Enduring Security Framework nsaesf@cyber.nsa.gov

Media Inquiries / Press Desk:

- NSA Media Relations, 443-634-0721, MediaRelations@nsa.gov
- CISA Media Relations, 703-235-2010, CISAMedia@cisa.dhs.gov

TABLE OF CONTENTS

BACKGROUND

The Enduring Security Framework (ESF) hosted a 5G study group comprised of government and industry experts over the course of eight weeks during the summer of 2020 to explore potential threat vectors and vulnerabilities inherent to 5G infrastructures. At the conclusion of the study, the group recommended a three-pronged approach to explore this threat space[1]:

1. Identify and assess threats posed to 5G;
2. Determine what standards and implementations can achieve a higher baseline of 5G security; and
3. Identify risks inherent to the cloud that affect 5G security.

In support of this task, the ESF established a 5G Cloud Working Panel to engage with experts across government and industry to document 5G cloud security challenges, threats, and potential mitigations, to include guidance, standards, and analytics. The result of this collaboration is a four-part series of publications that addresses the third task identified by the 5G study group: applying a threat-based approach to identify and mitigate risks in 5G networks that derives from the use of cloud technologies and providing mitigations that can be applied to harden 5G cloud infrastructures.

SCOPE

This four-part series builds on the ESF *Potential Threat Vectors to 5G Infrastructure* white paper, released in May 2021, which focused specifically on threats, vulnerabilities, and mitigations that apply to the deployment of 5G cloud infrastructures.[2]

Although all 5G network stakeholders can benefit from this guidance, the recommendations are intended for service providers and system integrators that build and configure 5G cloud infrastructures. This includes core network equipment vendors, cloud service providers, integrators, and mobile network operators. The audience for each set of recommendations will be identified throughout the series, providing a layered approach to building hardened 5G cloud deployments.

5G CLOUD SECURITY CHALLENGE OVERVIEW

5G networks are being designed to handle the bandwidth, compute, and storage requirements that will be required for a predicted massive increase in network capacity as

[1] The ESF is a cross-sector working group that operates under the auspices of Critical Infrastructure Partnership Advisory Council (CIPAC) to address threats and risks to the security and stability of U.S. national security systems. It is comprised of experts from the U.S. government as well as representatives from the Information Technology, Communications, and the Defense Industrial Base sectors. The ESF is charged with bringing together representatives from private and public sectors to work on intelligence-driven, shared cybersecurity challenges.

[2] ESF, *Potential Threat Vectors to 5G Infrastructure*, 2021. https://www.nsa.gov/news-features/press-room/Article/2601078/nsa-odni-and-cisa-release-5g-analysis-paper

well as connected devices. For scalability, resilience, and agility, 5G networks leverage cloud infrastructures in the radio access network, core, and network edge. Cloud technologies underpin the implementation of virtual networking in 5G, enabling the dynamic allocation and management of networks for specific use cases, mobile network operators, or customers.

A characteristic of cloud infrastructure that presents a significant security challenge in 5G is multitenancy, the use of a shared physical infrastructure by multiple cloud infrastructure customers, e.g., mobile network operators. Multitenancy highlights the need to harden and securely configure technologies that isolate the workloads (e.g., virtualization/ containerization) for each of those customers. In addition, cloud providers and mobile network operators may share security responsibilities in a manner that requires the operators to take responsibility to secure their tenancy "in the cloud." An additional factor creating security challenges is the increasing deployment of a multi-cloud deployment model in 5G with diverse and evolving architectures and design approaches used by wireless carriers.

5G THREAT

Among the threat vectors presented in the *Potential Threat Vectors to 5G Infrastructure* analysis paper, several pertained to 5G cloud infrastructure, including *Software/ Configuration*, *Network Security*, *Network Slicing*, and *Software Defined Networking*.

5G networks, which are cloud-native, will be a lucrative target for cyber threat actors who wish to deny or degrade network resources or otherwise compromise information. To counter this threat, it is imperative that 5G cloud infrastructures be built and configured securely, with capabilities in place to detect and respond to threats, providing a hardened environment for deploying secure network functions. It is also important that 5G network functions be implemented using security best practices. This four-part series will address the former, providing guidance on hardening 5G cloud infrastructure deployments that are driven by threat information. This approach supports the May 2021 Presidential Executive Order on *Improving the Nation's Cybersecurity*, which called for secure products and services and enabling easier detection of unexpected behaviors and actions.[3]

5G CLOUD SECURITY GUIDANCE

Based on preliminary analysis and threat assessment, the Cloud Working Panel concluded that the top 5G cloud infrastructure security challenges could be divided into a four-part series that addressed different aspects of securing 5G clouds, facilitating the application of broad sets of mitigations.

[3] Executive Office of the President, *Executive Order on Improving the Nation's Cybersecurity*, 2021. https://www.whitehouse.gov/briefing-room/presidential-actions/2021/05/12/executive-order-on-improving-the-nations-cybersecurity

- **Part I: Prevent and Detect Lateral Movement**: Detect malicious cyber actor activity in 5G clouds and prevent actors from leveraging the compromise of a single cloud resource to compromise the entire network.

- **Part II: Securely Isolate Network Resources**: Ensure that there is secure isolation among customer resources with emphasis on securing the container stack that supports the running of virtual network functions.

- **Part III: Protect Data in Transit, In-Use, and at Rest**: Ensure that network and customer data is secured during all phases of the data lifecycle (at-rest, in transit, while being processed, upon destruction).

- **Part IV: Ensure Integrity of Infrastructure**: Ensure that 5G cloud resources (e.g., container images, templates, configuration) are not modified without authorization.

Zero Trust is the concept that perimeter defenses are no longer sufficient to secure a network, and that there should always be an assumption that a threat actor has established a foothold in the network[4]. This four-part series will document best practices that strive to bring a Zero Trust mindset into 5G cloud endpoints and growing multi-cloud environments. All actions should be explicitly verified and monitored. Although the best practices documented in this series do not constitute a complete Zero Trust template for securing 5G cloud infrastructures, if the best practices are applied, a 5G cloud environment will have made significant strides toward the implementation of Zero Trust principles.

[4] NIST Special Publication 800-207. Zero Trust Architectures.
https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-207.pdf

DATA PROTECTION

A 5G Cloud Infrastructure comprises four security domains:

1. **Workload:** Virtual network functions (VNF) and cloud native network functions (CNF, previously referred to as Containerized Network Functions) deployed on virtual machines or containers, respectively.

2. **Platform:** Hardware, software, and network that supports workloads.

3. **Front-end Networks:** Network connectivity between the platform and other networks.

4. **Back-end Networks:** Network connectivity between the platform and Data Center Operations.[5, 6]

Part III focuses on protecting the confidentiality and integrity of data within a 5G cloud infrastructure. Data confidentiality measures should be designed to protect sensitive information from unauthorized access. Data integrity ensures that data is not tampered with or altered by unauthorized access. Authenticity mechanisms play a key role in data protection by confirming users and systems are authorized with the correct rights to access the 5G cloud infrastructure data. [7]

[5] CNTT Cloud iNfrastructure Telco Taskforce Reference Model.
https://cntt.readthedocs.io/en/stable-elbrus/ref_model/chapters/chapter01.html
[6] CNTT Cloud iNfrastructure Telco Taskforce Chapter 7.
https://cntt.readthedocs.io/en/stable-elbrus/ref_model/chapters/chapter07.html#7.4.1
[7] Ibid p4.

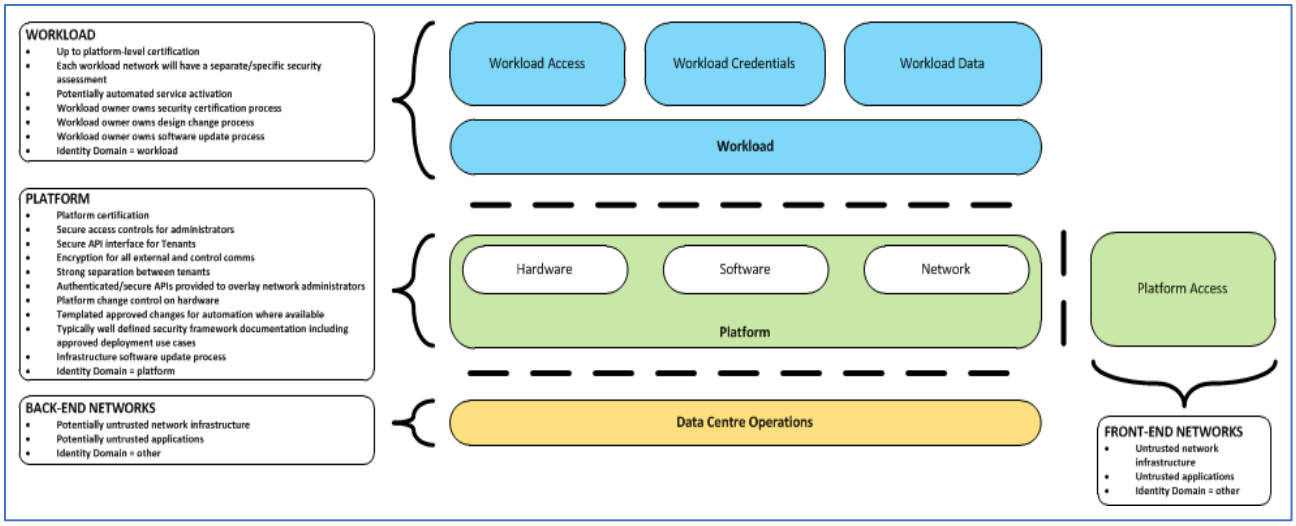

Figure 1: 5G cloud infrastructure security domains

CONFIDENTIALITY, INTEGRITY, AVAILABILITY (CIA) TRIAD

The confidentiality, integrity, and availability (CIA) triad drives the requirements for secure 5G cloud infrastructure systems and data. Figure 1 illustrates the 5G cloud infrastructure security domains and several high-level requirements for achieving CIA protection in each domain.

Audience: Cloud Providers, Mobile Network Operators, Customers

Guidance/Mitigations

- The Platform must support confidentiality and integrity of data at-rest, in-transit, as well as related metadata. [8]
- The Platform must support confidentiality and integrity of processes and restrict information sharing with only authorized parties (e.g., tenant).
- The Platform must support confidentiality and integrity of process-related metadata and restrict information sharing with only authorized parties (e.g., tenant)
- The Platform must support confidentiality and integrity of workload resource utilization (RAM, CPU, Storage, Network I/O, cache, hardware offload) and restrict information sharing with only authorized parties.
- The Platform must not allow memory inspection by any actor other than the authorized actors for the Entity to which Memory is assigned (e.g., tenants owning the workload), for Lawful Inspection, and by secure monitoring services.

[8] Ibid p4.

- The monitoring system must not affect the data confidentiality of the infrastructure, workloads, or the user data.

PROTECTION OF DATA-IN-TRANSIT

In a 5G context, data in transit applies in two different planes: the control plane (CP) and the user plane (UP). In 5G the control plane signaling data is encrypted via Transport Layer Security (TLS)[9]. Work is underway within the GSMA[10] to define the implementation specifics for how TLS is implemented in the control plane, and internet between networks. The algorithms required are specified by the 3GPP.[11]

Control plane data confidentiality and integrity are both required capabilities on 5G endpoint devices and 5G base stations[12]. All control plane data between the endpoint device and the base station (with a few exceptions, including unauthenticated emergency calls), must have integrity protection. However, confidentiality for control plane data remains optional.

User plane data confidentiality and integrity capabilities are required, but their use is optional at the discretion of the operator. This is optional primarily due to the additional processing load at the user equipment and base station and its impact to the size of the resulting communication packets.

Audience: Cloud Providers, Mobile Network Operators

Guidance/Mitigations

- Some of the user plane threats, such as Person-in-the-Middle and privacy violations, may be mitigated through the required use of the optional confidentiality and required integrity capabilities discussed above. Others, such as routing and Denial of Service (DoS) attacks must be handled in the control plane and above and would benefit from the required use of both the optional confidentiality and integrity capabilities discussed above.

- Where there are multiple hosting facilities used in the provisioning of a service, network communications between the facilities for the purpose of backup, management, and workload communications should be cryptographically protected in transit between data center facilities.[13]

[9] NIST Special Publication 800-52 – Guidelines for the Selection, Configuration, and Use of Transport Layer Security (TLS) Implementations. https://csrc.nist.gov/publications/detail/sp/800-52/rev-2/final
[10] GSMA Cloud Infrastructure Reference Model v1.0.
https://www.gsma.com/newsroom/wp-content/uploads//NG.126-v1.0-1.pdf
[11] 3GPP Portal Specification #: 33.501.
https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3169
[12] Ibid p4.
[13] Ibid p4.

- Systems transmitting data should use protocols that limit security risk such as SNMPv3, SSH v2, ICMP, NTP, syslog, and TLS v1.2 or higher. [14] For example, mutual authentication must be performed before encrypted data is sent from one system to another.

- Ensure all forms of data in transit are protected using strong cryptographic algorithms with strong integrity protection. Select cryptographic algorithms, modes and key sizes from the Commercial National Security Algorithm Suite (CNSA)[15] when applicable.

- These mitigations require the use of key and certificate management systems (preferably global or federated, rather than *ad hoc*) between organizations sending and receiving this encrypted data.

- Multiple cloud-based Hardware Security Modules (HSMs) should be employed where practical and should be required as a Root-of-Trust for high-risk or high-value data transmissions. This will also aid availability, data security monitoring, and governance.

PROTECTION OF DATA-AT-REST

5G data at rest is provided by any 5G network function responsible for storing data used in user plane and control plane processing.

Protecting data at rest in a 5G solution must meet 3GPP requirements in addition to meeting local regulations related to protecting sensitive and confidential data.

Data at rest in a cloud environment, specifically a 5G cloud infrastructure, can exist in multiple forms. Examples of forms of data at rest include:

- Persistent subscriber-level application data that allows subscribers to access the 5G network.

- Persistent data that affects and tracks 5G Network Function (NF) processing.

- Ephemeral data that affects and tracks 5G NF processing.

- Confidential system internal data that controls and defines the NF. Confidential system internal data includes authentication data (e.g. PINs, cryptographic keys, passwords, and cookies) as well as system internal data that is not required for

[14] Ibid p4.
[15] Commercial National Security Algorithm Suite.
https://apps.nsa.gov/iaarchive/programs/iad-initiatives/cnsa-suite.cfm

systems administrators and could be of advantage to attackers (e.g. error messages containing stack traces).[16]

Data at rest can reside in primary, replica or backup storage. All forms of storage related to data at rest must meet a minimum set of requirements for protecting data at rest. Guidance related to protecting data at rest is available from the National Institute of Standards and Technology (NIST).

5G subscriber data exists in both storage related to the profile of the subscriber in the 5G environment as well as in tracing or logging information related to the subscriber. Subscriber data contains Personally Identifiable Information (PII) that defines unique characteristics of a subscriber as well as sensitive subscriber data elements like the long-term key K.[17] PII and sensitive subscriber data must be protected when the data is at rest.

Confidential system internal data must be protected as well by ensuring that access is limited and the data at rest is protected by strong cryptography and access rules.

Audience: Cloud Providers, Mobile Network Operators, Customers

Guidance/Mitigations

- All data persisted to primary, replica, or backup storage is to be encrypted.[18]
- Ensure all forms of data at rest are protected using strong cryptographic algorithms with strong integrity protection. Select cryptographic algorithms, modes and key sizes from the Commercial National Security Algorithm Suite (CNSA)[15] when applicable.
- Refresh cryptographic keys, used to protect data, periodically. A good practice involves refreshing keys at least once a year.[19]
- Best practice is to secure the workload volumes by encrypting them and storing the cryptographic keys at multiple safe locations.
- The hypervisor should be configured to securely erase the virtual volume disks in the event of application crashes or is intentionally destroyed to prevent it from unauthorized access.
- Clean all subscriber data removed from the data at rest storage
- The Platform should support self-encrypting storage devices.

[16] 3GPP Portal Specification #: 33.117, Section 4.2.3.2.2. https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=2928
[17] Ibid p6.
[18] Ibid p4
[19] NIST Special Publication 800-53 Rev 5. Security and Privacy Controls for Information Systems and Organizations. https://csrc.nist.gov/publications/detail/sp/800-53/rev-5/final

- Institute policies and processes that evaluate and categorize data to ensure that data containing sensitive and confidential attributes receive the proper level of protection.

- Perform security-related testing and auditing of environments that store data at rest to ensure the effectiveness of the protection scheme and the protection of all sensitive and confidential data.

- Ensure that access, Identity and Access Management (IAM), to data at rest is secured in a manner that strictly controls access to the data at rest according to the role, or access needs, required by the accessor.

- Ensure that access to data is traceable by ensuring that all accessors of data are uniquely identifiable.

- Ensure the availability of the data by performing real-time or near real-time back-ups of the data in order to protect from attacks (e.g., Ransomware attacks) and facilitate recovery from successful attacks.

- User authentication related to access to data at rest should use multi-factor authentication or Public Key Infrastructure (PKI) based certificate authentication.

- Ensure that tools are in place to detect data integrity impacting events and processes exist that define recovery procedures.

- Cryptography enlisted to protect data at rest must be defined and approved by relevant standards bodies such as the NIST. NIST publications provide guidance related to the use of approved cryptographic functions.

- All forms of storage related to data at rest, such as primary, replica, and backup, must meet the minimum requirements in terms of securing the data.

- Backup storage can incorporate data integrity protection measures like a write once and read many approaches.

- The Platform must support Secure Provisioning, Availability, and Deprovisioning (Secure Clean-Up) of workload resources where Secure Clean-Up includes tear-down, defense against virus, or other attacks. Note: Secure clean-up: tear-down, defending against virus or other attacks, or observing of cryptographic or user service data.

- Like data-in-transit, multiple cloud-based Hardware Security Modules (HSMs) should be employed where practical and should be required as a Root-of-Trust for high-risk or high-value data-at-rest. This will also aid regulatory requirements, availability, data security monitoring, and governance.

PROTECTION OF DATA-IN-USE

It is universal practice currently in cloud and enterprise environments to protect data-at-rest using strong encryption in local and/or network-attached storage. However, when the same data is being processed by the central processing unit (CPU), it is held as plain text in memory and not protected by encryption. Memory contains high-value assets such as storage encryption keys, session keys, credentials, PII, customer IP, and important system data. Virtualization, cloud, and multi-tenancy brings the additional dimension where environments (ex. VMs or Containers) from two different customers could be running on the same machine. For sensitive / regulated workloads, there is a desire to protect data- in- use from the underlying privileged system stack as well as physical access threats. Therefore, it is critical that data in memory has comparable protection to data-at-rest. This is the focus of confidential computing – protecting data in use on compute devices using hardware-based techniques like Trusted Execution Environments.

Plaintext sensitive data elements, such as the long-term key K, should not leave the boundaries of the components that use the data elements. The Authentication credential Repository and Processing Function (ARPF) that resides in the 5G Unified Data Management (UDM) Network Function (NF) is an example of a boundary that uses sensitive subscriber data.

A Trusted Execution Environment (TEE) is an area in memory protected by the processor in a computing device. Hardware ensures confidentiality and integrity of code and data inside a TEE. The code that runs in the TEE is authorized, attested, and verified. Data inside a TEE cannot be read or modified from outside the TEE even by privilege system processes. Data is only visible while in the CPU caches during execution.

TEEs reduce the need to trust firmware and software layered on the system processing the workload. The trusted compute base (TCB), the hardware, firmware, and software components acting as the trusted system, is very small. In most TEEs, the TCB is the CPU (hardware and microcode) and the code defined by the owner. In some cases, the code is just a specific application; in others, it might be a purpose-built micro OS and the application. The CPU includes the TEE and the Rich Execution Environment (REE), allowing for decisions on where applications and data should be processed according to protection needs. The REE executes non-sensitive data, whereas the TEE can be programmed to execute encryption functions or the processing of sensitive applications for instance.

Known TEE vulnerabilities of data-in-use include vulnerabilities in TEE code and infrastructure, the possibility for opening side-channels, and exposure of data by design outside of the TEE, e.g., code in TEE sends data to external application in the clear or fails to use TEE intrinsic encryption features.

Audience: Cloud Providers, Mobile Network Operators

Guidance/Mitigations

- Implement Source Code Analysis of Code prior to load into TEE
- Perform regular updates/patching of Systems & Firmware for latest security fixes
- Leverage secure design guidance for code developed for TEE uses
- Verify and validate code before lode into TEE using cryptographic methods such as Signature or hash checking.
- Threats with the Mitigations provided by TEEs including:
    - Malicious/compromised admin CSP: A bad actor at the Cloud Service Provider (CSP) cannot access the TEE memory even with physical access to the server.
    - Malicious/compromised tenant of a hypervisor: A rogue app or compromised component of the system cannot access the TEE memory, even with a privilege escalation on the virtual machine manager (VMM).
    - Malicious/compromised tenant of a Container Engine/Environment: A rogue app, rogue container, or compromised component of the system cannot access the TEE memory, even with a privilege escalation on the OS.
- Malicious/compromised network: A rogue app or actor using a compromised network cannot access the data/IP inside the TEE.
- Compromised firmware/BIOS: Tampered BIOS or firmware will not be able to access the TEE memory.
- Physical Access Attacks at the edge: A bad actor targeting the edge compute nodes cannot access the TEE memory even with physical access to the system.
- Malicious/compromised admin at the Edge: A bad actor at the edge provider cannot access the TEE memory even with physical access to the system.
- Utilize devices, systems, and infrastructure that provides access to TEEs
- Verify the devices, systems, and infrastructure that provide the TEEs have been regularly patched/updated and are current
- Leverage TEEs for Sensitive and Regulated workloads and data protection

CONCLUSION

Part III of this series focused on protecting the confidentiality, integrity, and availability of data within a 5G cloud infrastructure. Implementing 5G mitigations, based on cybersecurity risks against data in transit, at rest, or in use, will ensure that only authorized services or functions have access to data within the network.