# SECURITY GUIDANCE FOR 5G CLOUD INFRASTRUCTURES

## Part I:
## Prevent and Detect Lateral Movement

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, both 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

## Prevent and Detect Lateral Movement in the 5G Cloud

Effective network cybersecurity practices involve defending the perimeter while recognizing that attackers sometimes successfully penetrate perimeter defenses, requiring the implementation of controls that detect and prevent adversarial activities on the network. This Zero Trust mindset places an emphasis on locking down internal resources along with an intensive logging and monitoring regime. Attackers commonly follow initial exploitation of a network with attempts to pivot laterally, taking advantage of misconfigurations, vulnerabilities, or other weaknesses in a network to gain a more extensive presence. Zero Trust practices (e.g., explicit and continuous authorization, extensive logging) help detect and prevent this category of adversarial activity.

5G cloud native deployments are susceptible to attacks at several points, including the exploitation of web vulnerabilities in the administration portals used by customers and network operators, in the 5G core via malicious or vulnerable applications and misconfigurations in the virtual networking stack or in the core/RAN clouds. Whatever initial position an attacker starts from, it is critical that controls be placed at all layers of the cloud to detect the adversarial presence and prevent the attacker from moving further.

Part I of this series presents guidance for mitigating lateral movement attempts by attackers who have successfully exploited a vulnerability to gain initial access into a 5G cloud system. Although this part focuses on a few critical areas, from a Zero Trust perspective, following the guidance provided in the other three parts of this series is equally important.

### Implement Secure Identity and Access Management (IdAM) in the 5G Cloud

After the initial compromise of a network, attackers commonly pivot laterally by exploiting the availability of internal services, particularly looking for services that are unauthenticated. For example, an attacker might use an initial position on a compromised virtual machine (VM) or container to access an application programming interface (API) or service endpoint that is not exposed externally. 5G cloud deployments will introduce more opportunities to move laterally in this manner because they support new implementations such as Service-Based Architecture (SBA), containers, and VMs that result in more element-to-element communications than in previous networks that utilized physical appliances and point-to-point interfaces. Reducing the risk of these types of attacks, both at the network-function layer as well as the underlying cloud infrastructure layer, is a critical activity for reducing the overall risk of lateral movement.

**Audience:** Cloud Providers, Mobile Network Operators

**Guidance/Mitigations**

- 5G networks should assign unique identities to all elements (and preferably to each interface) that will communicate to other elements in the 5G network.

- Before allowing access to a resource (e.g., Application Programming Interface (API), Command Line Interface (CLI)), each network element should authenticate and authorize the entity requesting access.

- Where possible, identities should be assigned using Public Key Infrastructure X.509 certificates from a trusted certificate authority (CA) rather than username/password combinations[5].

- If username/passwords must be used, multi-factor authentication (MFA) should be enabled to reduce the risk of compromise.

- The 5G network should provide automated mechanisms for credential management, especially as these features become more readily integrated in modern cloud environments (e.g., certificate rotation via a Service Mesh).

- Where possible, use certificate pinning or public key pinning to provide additional identity assurance when authentication is dependent upon multiple CAs. Certificate pinning and public key pinning associate a host with an expected certificate reducing the impact from a compromised CA.

- All access to resources should be logged. Each log entry should contain the time, resource, requesting entity (name or service), information about the requesting entity's location (region, IP address), and result of the access request (allow, deny). Logs should be protected as described in *Part III: Protect Data in Transit, In-Use, and at Rest* of this series.

- Analytics for detecting potentially malicious resource access attempts should be deployed and run regularly.

[5] Federal Public Key Infrastructure Guide Introduction. https://playbooks.idmanagement.gov/fpki/

### KEEP 5G CLOUD SOFTWARE UP-TO-DATE AND FREE FROM KNOWN VULNERABILITIES

5G cloud native deployments rely on the secure coordination of multiple services built from heterogeneous software sources. In addition to the basic services that make up a typical cloud, 5G clouds may deploy open source or specialized services to support network slicing, including third party applications that implement virtual network functionality. Vulnerabilities in any of this software could be exploited by an attacker to gain initial access into the 5G cloud infrastructure or enable an attacker, who has established a foothold in the cloud, to move laterally.

Software vulnerabilities fall into three categories: publicly known with a patch available from the software vendor; publicly known without a patch (n-day); and not publicly known (0-day). Although measures should be taken to mitigate the risk of n-day and 0-day vulnerabilities, patching publicly known vulnerabilities as quickly as possible significantly reduces the risk of exploitation. Maintaining the security of software used within 5G cloud environments is critically important to preventing adversarial lateral movement.

The guidance in this section applies to all software that runs in 5G cloud infrastructures, including the cloud/virtual networking software, as well the management and orchestration code for deploying virtual networks and any other integrated applications. All organizations that deploy software to the cloud have a responsibility to maintain secure software development practices.

**Audience**: Cloud Providers, Mobile Network Operators, Customers

**Guidance/Mitigations**

- Refer to *NIST Special Publication (SP) 800-40, Guide to Enterprise Patch Management Technologies.*

- Integrate source code scanning and patching into the software development and deployment process:

    - Regularly scan software repositories for known vulnerabilities and out-of-date versions using one or more software scanning tools or services;

    - Regularly monitor third party applications and libraries that are integrated into the network slicing infrastructure for publicly reported vulnerabilities;

    - Patch critical vulnerabilities in the operational environment within [policy-defined, suggested: < 15] days and other vulnerabilities within [policy-defined, suggested: <60] days.

### SECURELY CONFIGURE NETWORKING WITHIN THE 5G CLOUD

In a 5G cloud native deployment, two network functions, or microservices, may be sitting in the same logical network segment but may be members of two completely different security groups based on their functions. A security group may be constructed using a Network Access Control List (ACL) or a stateful firewall that determines which outbound or inbound connections are permitted, or a network slice instantiation. The same principle should also apply to underlying infrastructure. For example, a Kubernetes (K8s) cluster with Control Plane and Worker networks, should use networking functions (subnetting and stateful firewall/ACL) to control which nodes can communicate would add an additional layer of security.

**Audience**: Cloud Providers, Mobile Network Operators

### Guidance/Mitigations

- Create security groups per K8s Pod. Security groups for Pods make it easy to achieve network security compliance by running applications with varying network security requirements on shared compute resources.

- Use private networking for connecting microservices/network functions. This could be facilitated by a container networking plugin that enables attaching multiple network interfaces to Pods.

- Configure default firewall rules or default ACLs to block inbound and outbound connections at the Pod and worker node level. This is commonly provided by Kubernetes Network Policies. Some container networking interfaces[6] have enhanced filtering to protect the host on which K8s is deployed from Pod- and cluster-external traffic. [7]

- Use Service Meshes to protect node-to-node traffic. Service Meshes, in the context of networking, can provide end-to-end authentication and service monitoring through the use of "side cars", which are containers injected into each Pod. These side cars provide a proxy through which all TCP traffic is forced to traverse via Netfilters REDIRECT rules. This proxy can provide a mutual, TLS authenticated and encrypted connection between Pods in a cloud-native deployment. This protects Pods from external and Pod-to-Pod attacks.

## LOCK DOWN COMMUNICATIONS AMONG ISOLATED NETWORK FUNCTIONS

5G network implementations can have significantly more communication sessions between network elements than in 4G LTE. A network function (NF) can communicate over the control plane, user plane, management plane, and through the cloud infrastructure. The network elements may be isolated to meet the security requirements of the customer, network slice, or use case. Communications paths that rely on insecure authentication mechanisms, or that are not locked down sufficiently by policy, could be used as a lateral movement path by an attacker, allowing the attacker to change between planes or pivot to gain privileges on

[6] Container Network Interface (CNI) is a Cloud Native Computing Foundation project that provides networking for Linux containers. https://github.com/containernetworking

[7] CIS Benchmarks Securing Kubernetes, 2021. https://www.cisecurity.org/benchmark/kubernetes/

another set of isolated network resources. To effectively prevent and detect lateral movement, the 5G network must provide mechanisms to ensure that all such communication sessions are authorized, and policy is enforced over network resources in the same security group.

**Audience:** Cloud Providers, Mobile Network Operators

### Guidance/Mitigations

- 5G networks should ensure that all communication sessions on an NF's control plane, user plane, management plane, and through the cloud infrastructure are authenticated using the identities provisioned from the Identity and Authorization session. For example, these sessions could use mutually-authenticated TLS v1.2+ where the X.509 certificates are the identities that are authenticated.

- Policies should be created and deployed that enforce the separation of network resources in the same security group based on secure authentication and authorization.

### Monitor for Indications of Adversarial Lateral Movement

An attacker who has stolen legitimate, authorized user credentials or has exploited a vulnerability in the 5G cloud deployment and attempts to laterally move from within the mobile network operator's (MNO) network equipment will likely leave evidence of this movement. It is critical to continuously monitor for evidence of exploitation and adversarial lateral movement within 5G cloud deployments.[8]

**Audience:** Cloud Providers, Mobile Network Operators

### Guidance/Mitigations

- Suppose that a threat actor has obtained the legitimate credentials of an employee with access to the Operations, Administration, and Management (OA&M) section of a particular node within the 5G network cloud ecosystem. Most network equipment's OA&M interfaces should be on the protected OA&M internal protected network, whereby:
  - Users would have to gain access to an adjacent network that allows connection directly to the domain that the 5G network cloud OA&M equipment lives on;
  - Threat actors would have to gain access to the network first.

[8] FedRAMP Continuous Monitoring Strategy Guide
https://www.fedramp.gov/assets/resources/documents/CSP_Continuous_Monitoring_Strategy_Guide.pdf

- Most network equipment within the cloud is containerized; again, without credentialed access you cannot break from one container to another.

- A threat actor would then have to compromise several security protections in place, but assuming they do, there are several indicators that can be monitored to help detect this type of malicious activity:
    - Scanning behaviors or the opening of unusual ports between one network node's OA&M interfaces and another node's OA&M interfaces;
    - User behavior abnormalities (time of day usage, type of activity usage);

- Network communication abnormalities.
    - To execute exploits or extract information, the network may begin communicating in unusual ways, not necessarily to known "bad IPs," but to internal systems with which the network does not normally communicate.

- Pod/container logging abnormalities, such as unexpected system calls.
    - Container breakouts often rely on execution that is atypical for the containerized application. This anomalous behavior can be identified through comparison against a baseline of Pod behavior or through machine learning or AI-enabled security auditing.

### DEVELOP AND DEPLOY ANALYTICS TO DETECT SOPHISTICATED ADVERSARIAL PRESENCE

Detecting the presence of attackers or other security incidents within 5G cloud native deployments is challenging due to the massive amounts of network traffic and IdAM events occurring regularly. Sophisticated analytics, based on machine learning and artificial intelligence, can help detect adversarial activity within the cloud and provide the 5G stakeholder with the means to detect malicious use of customer cloud resources (e.g., networks, accounts).

The effectiveness of each analytic at detecting threat depends on the layer within the cloud as well as the 5G stakeholder that is deploying the analytic. 5G cloud providers may deploy analytics to detect low level attacks, such as hypervisor or container break-outs. MNOs may deploy analytics to detect the malicious use of network credentials. Each stakeholder must understand its role in threat detection and incident response.

Balancing data confidentiality requirements with the ability to inspect network traffic for threats is a challenging problem when developing analytics. The effectiveness of "break and inspect" and other techniques that require exposure to unencrypted network traffic must be weighed against privacy concerns and legal requirements. Analytics deployed at all layers must make risk decisions regarding analytic data requirements.

**Audience:** Cloud Providers, Mobile Network Operators

**Guidance/Mitigations**

- Stakeholders at all layers of the 5G cloud stack should leverage an analytic platform to develop and deploy analytics that process relevant data (cloud logs and other telemetry) available at that layer. The analytics should be capable of detecting known and anticipated threat, but also be designed to identify anomalies in the data that could indicate unanticipated threat.

## Conclusion

An attacker can use cloud/virtual networking to move through a network after initial compromise. Configurations to prevent and detect lateral movements is only one aspect of hardening a 5G cloud infrastructure. The detection and mitigation of lateral movement attempts within a 5G cloud system is a shared responsibility among 5G cloud providers, network operators, mobile network operators and customers.

In *Part II: Securely Isolate Network Resources,* the following topics will be addressed:

- Enabling isolation of Kubernetes Pod resources, such as limiting capabilities and permissions on deployed containers.

- Cryptographically isolating critical containers from infrastructure/hosts using trusted execution environments.

- Applying good container security hygiene to avoid resource contention and Denial of Service attacks.

- Implementing real-time threat detection and incident response through minimizing noise, curating baseline behavior, and alerting on anomalous activity.