Enhanced 5G and Zero Trust Cloud and Operational Security Aspects

TABLE OF CONTENTS

TABLE
OF
CONTENTS

EXECUTIVE SUMMARY

Zero Trust (ZT) is a concept that no digital system or human user, whether external or internal, can be trusted, regardless of ownership and location. Zero Trust Architecture (ZTA) is a plan to implement ZT in a digital system or network of digital systems. ATIS published a paper in July 2023, “Enhanced Zero Trust and 5G,” that focused on broadly implementing a ZTA in 5G. In it, ATIS enumerated 12 required central security controls for a 5G ZTA, along with numerous recommendations to 3rd Generation Partnership Project (3GPP) to align 5G and 6G security specifications with a ZTA.

5G Mobile Network Operators (MNOs) architect and deploy their 5G wireless offerings on fully virtualized or cloud-native platforms and networks. When implementing cloud-native networks, the cloud computing platforms host all the various parts of a 5G System, including the 5G Core Network, Operations and Business Support Systems (OSS and BSS), and Open Radio Access Networks (O-RAN).

At present, there are four dominant cloud models in production use. Two of these models are considered legacy in that they use traditional virtualized and cloud compute architectures, while the other two leverage the cloud services delivered by Hyperscaler Cloud Providers (HCPs). The HCP-based models are gaining momentum in the industry and driving new investments in the public cloud. The two legacy models are the multi-vendor stack and the single vendor full stack, which are private cloud implementations. The two public cloud models leverage well-known public cloud offerings from the cloud providers. One is the standard public offering, and the second is private and uses the same technology stacks but is dedicated specifically for the 5G MNO installed on their physical locations.

5G networks are currently being deployed in a hybrid cloud environment using combinations of the four models of private and public clouds. The 5G MNO must rely on several different Standards Development Organizations (SDOs) for cloud computing architecture and security.

This paper looks at the implementation and operational aspects of implementing a ZTA in the 5G MNO cloud environments that are hosting their 5G services. We will look at the four cloud deployment models in terms of a ZTA in combination with a discussion of the 12 security controls. We will highlight any potential deployment issues relating to these 12 security controls for each of the four models. The development and rollout of the security controls have also been primarily centered in the IT world. For some, there are gaps due to differences from the IT side that must be filled in order to deploy the controls in the 5G realm. This paper highlights these gaps, including threat intelligence feeds. Gaps filled by new threat models such as MITRE’s FiGHT and Groupe Spéciale Mobile Association (GSMA) Mobile Threat Intelligence Framework (MOTIF) are also highlighted.

The ATIS Cloud 5G ZTA study was informed by the work at U.S. Department of Commerce National Institute of Standards and Technology (NIST) and ZT subject matter experts from organizations that are stakeholders in 5G network security. The recommendations enumerated in Section 5 provide concluding strategic guidance to enhance security and operational resilience in 5G cloud environments through a ZT framework. Recommendations are broken down into sections for each responsible standards bodies, infrastructure vendors, HCPs, and security operations teams.

1. IMPLICATIONS AND SCOPE OF ZTA FOR 5G CLOUD

Traditionally, network infrastructure has been deployed within the MNO's premises where the MNO owns, manages, and controls all assets, including facilities, transport infrastructure, network services, network applications, and data. Mobile networks are evolving toward cloud-based architectures that introduce a layered implementation where multiple stakeholders, including vendors, integrators, and providers, may own, manage and control facilities, network services, network applications, and/or infrastructure.

The introduction of third parties and cloud-native deployments facilitate new threats from internal and external attacks on 5G critical infrastructure. External and internal threat actors could gain access to 5G virtual or cloud-native functions through the infrastructure, owned or managed by these third parties, to perform confidentiality, integrity, and availability attacks on the network. Within the cybersecurity domain, the act of a threat actor (internal and external) gaining a foothold in a network is called an Advanced Persistent Threat (APT). Most recently, Salt Typhoon is an APT that received considerable industry recognition [1]. In addition, many cloud models can obscure visibility into lower layers of the infrastructure, making it much more difficult to monitor for cybersecurity vulnerabilities and attacks. The impact of these attacks on 5G voice and data services could include outages, performance degradation, unauthorized reconnaissance, and data theft.

Mitigating these threats can be challenging because each of the third-party vendors/providers may have their own set of security controls that may or may not be sufficient and may not seamlessly interwork with other system-wide security controls. As such, ATIS has created a set of fundamental zero trust security control groups to help address this issue.

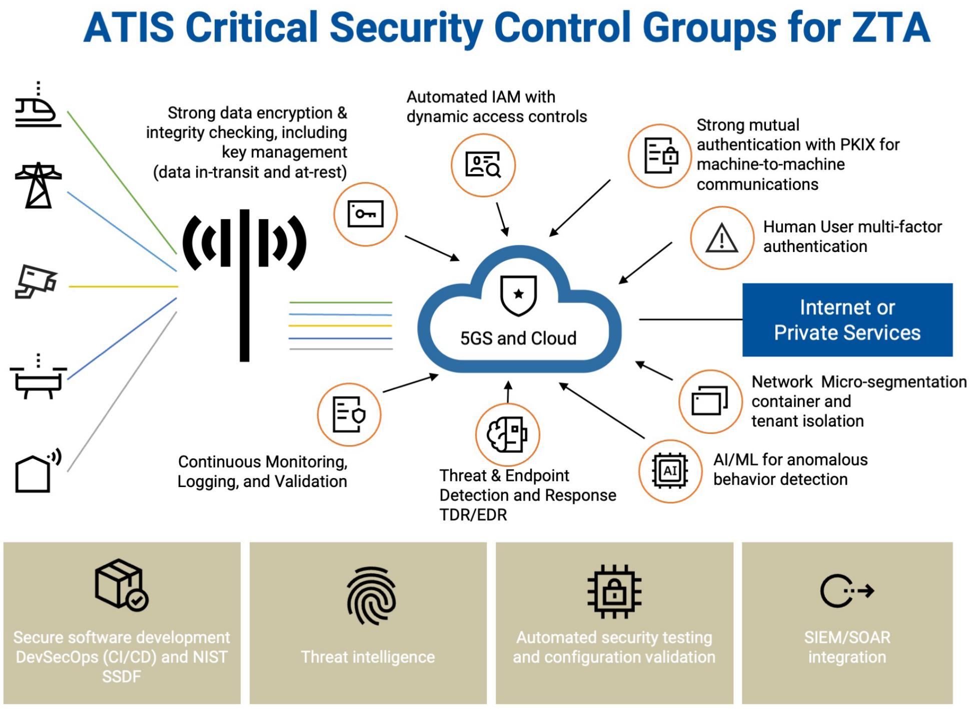

ATIS Critical Security Control Groups for ZTA
Strong data encryption & integrity checking, including key management (data in-transit and at-rest)
Automated IAM with dynamic access controls
Strong mutual authentication with PKIX for machine-to-machine communications
Human User multi-factor authentication
5GS and Cloud
Internet or Private Services
Network Micro-segmentation container and tenant isolation
AI/ML for anomalous behavior detection
Threat & Endpoint Detection and Response TDR/EDR
Continuous Monitoring, Logging, and Validation
Secure software development DevSecOps (CI/CD) and NIST SSDF
Threat intelligence
Automated security testing and configuration validation
SIEM/SOAR integration

Figure 1: ATIS' 12 ZTA Security Control Groups

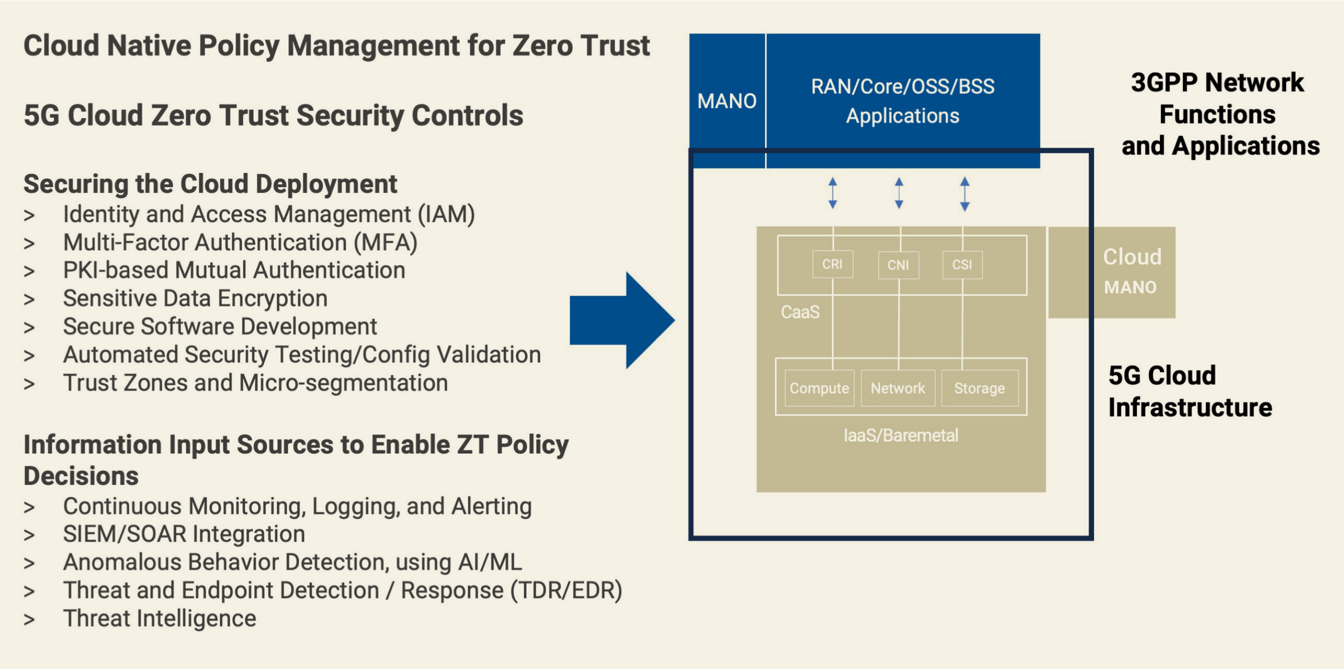
Figure 1 illustrates the 12 fundamental security control groups identified in the 2023 ATIS paper, "Enhanced Zero Trust and 5G" [2], to achieve a ZTA for 5G networks. These are:

1. Sensitive Data Encryption for data in motion, data at rest, and data in use
2. Identity and Access Management (IAM), including dynamic access control policies and the principle of least privilege
3. Public Key Infrastructure (PKI)-based Mutual Authentication for machine-to-machine communications
4. Multi-Factor Authentication (MFA) for human users
5. Network micro-segmentation and micro-perimeters
6. Anomalous Behavior Detection, using artificial intelligence/machine learning (AI/ML)
7. Threat and Endpoint Detection and Response (TDR/EDR)
8. Continuous Monitoring, Logging, and Alerting
9. Secure software development based upon the DevSecOps including continuous integration/ continuous deployment (CI/CD)[3] and NIST Secure Software Development Framework (SSDF)
10. Threat Intelligence
11. Automated Security Testing/Configuration Validation
12. Security Information Event Management (SIEM)/ Security Orchestration, Automation, and Response (SOAR) integration

The focus of this paper is to provide recommendations about how to implement a ZTA and all of the ATIS-defined security controls into the 5G cloud infrastructure. This leaves a gap in terms of what security controls are required within the infrastructure to define ZTA for the 5G cloud. One of the primary goals of this paper is to recommend how the ATIS-defined 12 critical security controls are implemented uniformly in the 5G cloud infrastructure to provide layers of defense required for a Defense in Depth strategy.

There is an overlap with some of these controls across the 3GPP Network Functions (NFs), applications, and management within the 5G cloud. Examples of this include security and audit logs, alerts, and telemetry being sourced from the 3GPP NFs and the applications, which are then streamed toward continuous monitoring and control systems such as the Security Information Event Management (SIEM) and EDR functions. SIEM and EDR agents may exist within the NF and/or the underlying cloud infrastructure.

Another key mechanism is an IAM system that controls human access to the NFs and applications. That same IAM can control access for the 5G cloud infrastructure components. This paper will give guidance about how these security controls should be deployed in the 5G cloud and the overlap into the 3GPP components. Gaps that exist along with recommendations to close them are enumerated for the appropriate standards or industry bodies.



Cloud Native Policy Management for Zero Trust

5G Cloud Zero Trust Security Controls

Securing the Cloud Deployment

> Identity and Access Management (IAM)
> Multi-Factor Authentication (MFA)
> PKI-based Mutual Authentication
> Sensitive Data Encryption
> Secure Software Development
> Automated Security Testing/Config Validation
> Trust Zones and Micro-segmentation

Information Input Sources to Enable ZT Policy Decisions

> Continuous Monitoring, Logging, and Alerting
> SIEM/SOAR Integration
> Anomalous Behavior Detection, using AI/ML
> Threat and Endpoint Detection / Response (TDR/EDR)
> Threat Intelligence

Figure 2: 12 ZT Security Control Groups Categorized

Figure 2 shows how the ATIS 12 ZT security control groups have been divided into two categories: "Securing the Cloud Environment" and "Information Input Sources to Enable ZT Policy Decisions." Given the critical role of policy management, we also include a section describing the role of policy management in ZT cloud deployments and the relationship to the security controls. Security controls should be deployed at each layer (Application, CaaS, Infrastructure) and visibility (via logging, Fault, Configuration, Accounting, Performance, and Security (FCAPS), EDR, and other telemetry) should be handled for each layer at that layer. In other words, one layer should not provide visibility at another layer. Telemetry across the layers can be sent to a centralized system, such as SIEM or Service Management and Orchestration (SMO), for centralized collection and correlation.

The focus of this document specifically addresses the intersection of ZT and 5G infrastructure (i.e., the 5G Core and Radio Access Network (RAN)) as deployed in cloud-native environments. ZT mechanisms for the operations and management environment of the 5G infrastructure are not specifically addressed in this document. In addition, important considerations such as the creation of a security architecture and associated trust boundaries are not addressed.

2. CURRENT STATE OF ZERO TRUST FOR 5G

2.1 NIST ZTA

NIST defines zero trust as a concept that no digital system or human user, whether external it internal, can be trusted, regardless of ownership and location. Zero Trust Architecture (ZTA) is a plan to implement ZT in a digital system or network of digital systems. However, NIST points out that, within a ZTA, an "implicit trust zone" can exist in a system which represents an area where all the entities are trusted to at least the level of the last PDP/PEP gateway. NIST advises that the implicit trust zone within the ZTA must be as small as possible [4].

NIST defined seven tenets as foundational for its ZT architecture:

List 1. NIST Seven Tenets of Zero Trust [4]

T1. All data sources and computing services are considered resources

T2. All communication is secured regardless of network location

T3. Access to individual resources is granted on a per-session basis

T4. Access to resources is determined by dynamic policy

T5. The operator monitors and measures the integrity and security posture of all owned and associated assets

T6. All resource authentication and authorization are dynamic and strictly enforced before access is allowed

T7. The operator collects information about the current state of assets, network infrastructure and communications and uses it to improve its security posture

The tenets are intended to provide a set of principles that are intentionally broad but also fully inclusive so that any gaps can be avoided.

2.2 Cybersecurity and Infrastructure Security Agency (CISA) Zero Trust Maturity Model (ZTMM)

The ZTMM is a framework that helps organizations assess and improve their ZTA implementation through a journey of four stages that provide a structured approach to measure progress and identify areas for improvement:

1. **Traditional Stage**

   Security controls are implemented only at the perimeter to protect against external threats.

2. **Initial Stage**

   Foundational security controls are in place and basic ZT principles are applied to protect NFs as micro-perimeters from external and internal threats. This consists of increased visibility and control of network traffic, and authentication and authorization of external and internal subjects requesting access to resources.

3. **Advanced Stage**

   More advanced security controls are in place to enable greater visibility and control over network traffic on external and internal interfaces. Micro-segmentation is utilized to prevent lateral movement within the network and the implementation of continuous authentication and authorization.

4. **Optimal Stage**

   AI/ML is leveraged for advanced threat detection and decision-making. This allows for AI/ML driven automated responses to identified threats and for continuous, automatic optimization of security policies.

2.3 3GPP

3GPP has undertaken a phased approach to the evaluation of 3GPP architecture for all seven of the NIST ZT tenets. The initial focus has been on evaluation of the 3GPP 5G Core Service-Based Architecture (SBA), while the O-RAN Alliance is focused on ZTA for Open RAN.

The first 3GPP evaluation was undertaken and documented in 3GPP Release 18 Technical Report (TR) 33.894 [5], which noted that Tenets 1, 2, and 3 required no further study as applicable security mechanisms already exists. For the remaining Tenets, although this report did not agree to standardize solutions for ZT, it did allow for further studies related to exposing information for security monitoring and policy enforcement. A subsequent evaluation of the 3GPP 5G Core SBA was undertaken in 3GPP Release 19 TR 33.794 [6], where the focus was on two aspects: 1) data exposure for security evaluation and monitoring and 2) security mechanism for dynamic policy enforcement. For data exposure to enable security evaluation and monitoring, it was concluded that it can be addressed with NF requirements to generate security event logs so data can be collected at the SBA layer for the following security incidents/scenarios:

1) Authentication and authorization failure event
2) Unexpected setup of Transport Layer Security (TLS) session and Application Programming Interface (API) invocation related to unauthorized reconnaissance
3) Malformed message event
4) High service load
5) Unexpected SBI call flows
6) Unexpected use of APIs exposed by services in SBA layer

As of this writing, further progress was made, and 3GPP approved a Rel. 20 work item for 5G-Advanced features on security-related events handling [7]. The main objective of this work item is to specify the data collection requirements that can be used for security purposes, which includes the following:

> General requirements for security events handling and collection

> Security requirements to transfer or communicate security events

> To specify events that need to be reported

Further work is needed for policy enforcement in the 5G Advanced and 6G systems. ATIS is working on a subsequent publication that will provide additional industry insight into what is required.

2.4
ETSI NFV-SEC Expert Group

The Security Expert Group of ETSI Network Function Virtualization (NFV) (ETSI NFV-SEC) is responsible for developing specifications around Whole System Security Management and Monitoring including:

> Virtualized Network Function (VNF) Package Security

> VNF Lifecycle Management Security

> Dynamic Certificate Management for Virtual Entities

> Isolation and Trust Domains (Containers)

> API Access (Tokens)

> And more

ETSI GS NFV-SEC024 Security Management Specification [8] provides a number of ZT-related specifications such as those relating to component compromise, multi-layer security, ZT security as an overlay, trust-but-verify approach, resiliency and redundancy, and more. This work was delayed during the COVID-19 pandemic, but ETSI NFV-SEC published updated drafts in October 2023 October 2024. Their plan is to finalize their draft by end of May 2025 and publish the final document in August 2025.

In March 2024, ETSI MEC Working Group published a Study on MEC Security [9] that references 3GPP's TR 33.894 [5] ZT study as something that they are tracking. ETSI is studying the security aspects related to Multi-access Edge Compute (MEC) application provenance verification with cryptographic attestations, including leveraging secure boot processes that incorporate a hardware root of trust and aspects of Internet Engineering Task Force's (IETF) Remote ATtestation procedureS (RATS) Architecture (RFC 9334) [10]. This study has identified key issues relating to stolen MEC application access tokens, stolen MEC application identity, compromised MEC applications and asset theft, compromise of application package during on-boarding, compromise of application during updates, threats associated with application package deletion, and MEC application anomalous behavior.

2.5
O-RAN ALLIANCE

O-RAN ALLIANCE [11] is committed to the security of the entire O-RAN architecture. O-RAN security is led by the O-RAN ALLIANCE's Security Work Group 11 (WG11), which is responsible for threat and risk analysis and normative security specifications. One of the approaches to open the RAN ecosystem is to have cloud-native NFs run on O-RAN-optimized cloud infrastructure, which is referred to as the O-Cloud. O-Cloud is specified by the O-RAN ALLIANCE WG6, with security requirements specified by WG11.

The WG11 technical report "O-RAN Study on Security for O-Cloud" [12] analyzes the security of O-Cloud architectural components and interfaces and provides a risk assessment of O-Cloud hardware and software infrastructures managed by the Service Management and Orchestration (SMO) via the O2 interface. The O-RAN Security Requirements and Controls Specification provides the normative security requirements for O-Cloud, SMO, and the O2 interface, along with the other O-RAN architecture elements and interfaces.

O-RAN Alliance is pursuing a ZTA [13] with core principles of continuous monitoring, authentication and authorization for internal and external subjects requesting access to O-RAN resources, least privilege access, and confidentiality and integrity protection of data, plus other comprehensive security controls throughout the O-RAN architecture, to protect against continuously evolving threats. The O-RAN Alliance is following the CISA ZTMM [14] to reach a ZTA in a phased approach.

O-RAN Alliance has identified O-RAN security specifications as being at the Initial stage of the CISA ZTMM, with plans to progress to the Advanced stage and eventually the Optimized stage[14]. To guide progress through the CISA ZTMM stages, WG11 has a ZTA work item. The work item's objectives are to study the applicability of the seven ZT tenets defined in NIST Special Publication (SP) 800-207 "Zero Trust Architecture"[4] to O-RAN Alliance's current security architecture and specifications, conduct an analysis on the current alignment between the seven tenets and O-RAN's security requirements and controls, and define new specifications where needed.

In October 2024, WG 11 published version 1 of their technical report for this work item, "Study on Zero Trust Architecture for O-RAN" [15]. In this technical report, WG 11 determined that at varying level of degrees, all seven tenets are applicable to O-RAN Alliance's architecture. Fully realizing all seven tenets may require integration with systems outside of the scope of O-RAN Alliance. This will have to be left up to each MNO implementation. For its part, O-RAN Alliance will specify security requirements, controls, and specifications that support the NIST tenets. WG 11 is still in the process of reviewing the current O-RAN architecture, including O-Cloud, to develop these new security requirements, controls, and specifications as needed.

2.6 Federal Communications Commission (FCC) CSRIC

Guided by the U.S. National Cybersecurity Strategy published on March 2, 2023 [16], the CSRIC VIII Report on Recommendations on the Role of the FCC in Promoting the Availability of Standards for 5G Environment of Virtualization Technology [17], responds to the four inquiries presented to the Communications Security, Reliability, and Interoperability Council (CSRIC) VIII by the FCC.

- **Query 1**: "Steps that the FCC should take (if any) to help coordinate formal standards, informal standards, and any collaborative open-interface community efforts to ensure interoperability in the virtualized 5G space."
- **Query 2**: "Recommendations on how the FCC can promote collaborations to achieve innovation in virtualized 5G."
- **Query 3**: "Recommendations on actions the FCC can take to build confidence in virtualized 5G solutions based upon open-source cloud computing software."
- **Query 4**: "Any other ways in which FCC can promote a diverse, competitive 5G environment."

The CSRIC report, which was developed through a cross-private and public sector collaboration effort, goes on to provide detailed recommendations for each of these queries. Most relevant to this document is the committee's recommendations on security and resiliency in response to Query 4.

The report recommends that the FCC can build confidence in virtualized 5G solutions by ensuring that security is built into systems from outset, rather than security being treated as something to be added later. The FCC should work with other government agencies to develop concise security requirements for virtualized 5G telecommunication environments based on ZT. Furthermore, the FCC should work with relevant U.S. government agencies to support the development of open-source test and evaluation tools to test these requirements.

3. 5G CLOUD DEPLOYMENT MODELS

3.1 5G Cloud Infrastructure

5G cloud infrastructure is the set of computation, storage, and networking equipment with related software that offers physical and/or virtual cloud resources and services to host 5G NFs and management systems [18].

Examples of 5G NFs and management systems hosted on 5G cloud infrastructures include the 5G Core, O-RAN NFs (i.e., O-CU-CP, O-CU-UP, and O-DU), and O-RAN SMO functions [19].

Table 1 describes the typical resources for a 5G cloud infrastructure, which requires management, orchestration, and workflow management services to include, but are not limited to, the following functionalities:

> Discovery and administration of 5G cloud infrastructure resources (i.e., storage, compute, networking)

> Software management of 5G cloud infrastructure, NFs, and management systems [19]

> Scale-In, Scale-Out of 5G cloud infrastructure resources, NFs, and management systems

> FCAPS (PM, CM, FM, Communication Surveillance) of 5G cloud infrastructure resources, NFs, and management systems

5G Cloud Infrastructure Resources	Description
Operating System	In virtualized host and guest environments, the OS runs within each VM. Alternatively, if there is no underlying hypervisor present, the operating system runs directly on the bare metal hardware.
Virtual Machine	A virtual guest environment executed on a hypervisor on a host. A set of system isolation technologies that provide various degrees of security isolation with the host machine's OS kernel.
Containers	A method for packaging and securely running an application that allows execution of multiple isolated user space instances while sharing the same underlying OS kernel.
Virtual Network Infrastructure	For communications within and between virtual machines and containers.
Hypervisor	When virtualization is used to manage resources, the hypervisor is responsible for allocating resources to each virtual machine. It may also be leveraged for implementing security.
Processing & Memory	The physical hardware that supplies central processing unit (CPU) time and physical memory.
Data Storage	The physical hardware used for file storage.
Network	This can be a physical or virtual network. It is responsible for carrying communications between systems and possibly the internet.
Physical Facility	The actual physical building where the cloud systems are located.

Table 1: 5G Cloud Infrastructure Resources
(Adapted from PCI SSC Cloud Computing Guidelines [20])

NIST SP 500-292 NIST Cloud Computing Reference Architecture [21] breaks down the three major types of Cloud Service Models. For 5G MNOs (the cloud consumer in the NIST models), these service types are relevant to the public cloud and related private cloud offerings from the various cloud providers.

> **Infrastructure as a Service (IaaS)**: The cloud provider provides the physical infrastructure (data centers, servers, hypervisors, network). The 5G MNO is responsible for the operating systems, virtual machines, container software systems, and containers that run the 5G NFs.

> **Platform as a Service (PaaS):** The cloud provider provides the platform, which includes the operating systems, virtual machines, and the Container-as-a-Service (CaaS) software components. The CaaS is located in the middleware layer defined by NIST. 5G MNOs load their NFs as Linux containers and virtual machines. Everything below the NFs themselves is the responsibility of the cloud provider PaaS service.

> **Software as a Service (SaaS)**: The cloud provider provides the infrastructure and platform. The application can be owned by an application vendor, which can be different from the cloud provider. The 5G MNO purchases the use of that application and is responsible for securing user access and managing data.

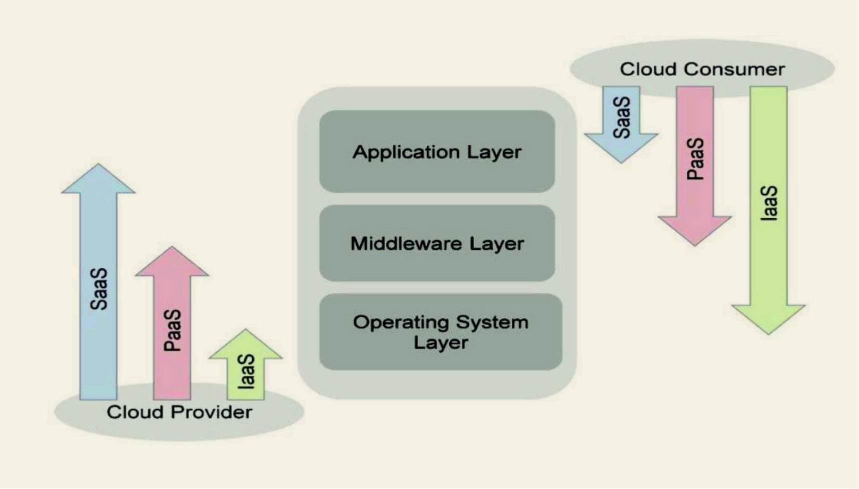


Figure 3: NIST SP 500-292 NIST Cloud Computing Reference Architecture [21]

3.2
Four 5G Cloud Deployment Models

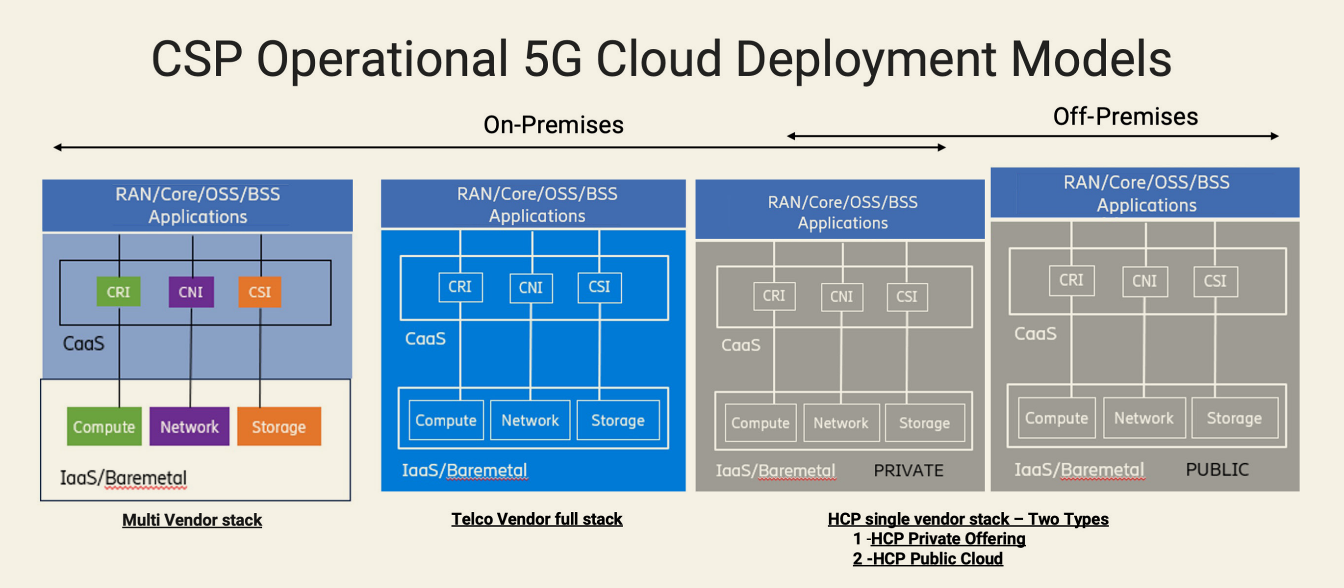


Figure 4: Four 5G Cloud Deployment Models

Within 5G MNO cloud networks, there are four concurrent cloud models that are in production use. Two of these are legacy, and two of them constitute a movement into the public cloud, where much of the present cloud investment is moving. The two legacy models are the multi-vendor stack and the telco vendor full stack.

In prior years, 5G MNOs built private clouds in their data centers using these two models, which support both Linux-based VNFs and Container Network Functions (CNFs) for their telco applications. With 5G, much of the investment is being transferred into two variations of the offerings with the Hyperscale Cloud Providers (HCPs). These are public cloud and a variant of the public cloud, which is an HCP private offering that uses the same technology stacks, but is dedicated specifically for the 5G MNO on their physical location, in a configuration also referred to as Hybrid Cloud.

At present, it is common to find all four models currently in both proof-of-concept testing and production. This current state constitutes a hybrid cloud environment using combinations of these private and public cloud models.

3.2.1 Multi-Vendor Stack

The multi-vendor stack is composed of different vendors for the server hardware, virtual or cloud-native NFs, and data storage. The CaaS portion hosts the 5G NFs, which run as containers provided by the telco infrastructure vendor. Each 5G MNO has historically chosen its own mix of vendors for their multi-vendor stacks, which causes a large variety of combinations across different 5G MNOs. In this model, the 5G MNO has complete control over all layers and components of the cloud stack, and they function as the primary integrator responsible for making all of the component parts work together from an operational standpoint. This model is deployed on the physical premises of the 5G MNO.

3.2.2 Telco Vendor Full Stack

In the vendor full stack model, the infrastructure vendor provides both the various NFs that run the mobile network and the full cloud-based infrastructure including the server hardware, the applications, and the data storage. The infrastructure vendor selects the different vendors that will make up the cloud stack layers. The CaaS architecture is chosen by the infrastructure vendor and can be built around open-source solutions that are part of the Linux distributions used. The infrastructure vendor is responsible for the integration of all of these various components and for the operational health of the overall system. Due to the responsibility belonging to the infrastructure vendor, this solution has historically been widely deployed in the 5G MNO data centers, along with implementations of the multi-vendor stack model. This model is deployed on the physical premises of the 5G MNO.

3.2.3 HCP Vendor Public Offering

This model is the well-known public cloud offering of the various HCPs, with the market being dominated by the big three: Amazon Web Services, Microsoft Azure, and Google Cloud. Here the infrastructure vendor's 5G NFs are deployed in the HCP's standard public cloud offering and run in the HCP's data center. The HCPs utilize their standard infrastructure offering of server hardware, network, and data storage. The model has the benefit of the HCP's owning the infrastructure and bearing the operational integration costs for the cloud infrastructure and is thus economically attractive.

Container as a service is a standard offering from the HCP. The 5G MNOs have been deploying applications into the public cloud now for some time. In particular, IT-related applications and Operations and Business Support Systems (OSS/BSS) are popular choices for deployment. The HCP controls the cloud infrastructure. The shared responsibility model assigns the HCP control of the cloud infrastructure and MNO control of the NFs and applications, along with responsibility to properly configure and operate those workloads running on the HCP's cloud infrastructure [22],[23].

3.2.4 HCP Vendor Private Cloud

Each of the big three HCPs has a private offering that utilizes the same technology and equipment that is used in the public cloud. However, it is deployed on the physical premises of the 5G MNO. Examples of such offerings are Amazon Web Service Outpost, Microsoft Azure for Operators Distributed Services, and Google Distributed Cloud Edge.

The private offering keeps the 5G MNO NFs, applications, and data on its premises and allows greater flexibility in the way that the cloud stack is deployed. This is important because 5G requires different levels of reliability and resiliency, along with different latency constraints, and these differ from the public cloud offerings. Because of this flexibility, 5G MNOs look to this model as a favorite to deploy the RAN and core NFs. The shared responsibility model assigns the HCP control of the cloud infrastructure and MNO control of the CaaS layer and applications running on top of it.

3.3
Implications of the 5G Cloud Deployment Models

All four models are currently deployed in some manner across the various 5G MNOs. This must be taken into consideration when deploying a security architecture with the security controls required to secure both the telco applications, NFs, and the cloud infrastructure itself. ATIS has defined 12 security controls that are required for a 5G ZTA. The effectiveness and scope of some of these controls depends on access to various levels of the cloud infrastructure; one example is logging sources. The various models offer different levels of logging for security visibility depending on the level of access allowed to the logs at different layers in the infrastructure. This will affect the amount of visibility offered for continuous monitoring.

Another example of this is the EDR agent. These agents require visibility into Linux system calls, which are requests by an application to the operating system to do a task such as open a file or to access a hardware device. In a Linux 5G environment using containers, the NF applications run in containers in user mode, which are separate from the underlying operating system, which runs in the Linux kernel. EDR agent implementations vary, with some running in user mode and others in both user and kernel mode.

Kernel mode access provides more visibility to system calls and internal information, and there is a debate concerning user versus kernel mode EDR agent effectiveness and scope of coverage. Access to the operating system, which runs in the kernel, varies by the cloud model in use. Those selecting a user and kernel mode EDR may be denied access to the operating systems in some cloud models, such as the public and private HCP cloud offerings. Such nuances concerning the security controls will be discussed in the related sections of this paper.

Considerations should be made for user versus kernel mode EDR agent, including the effectiveness and scope of coverage. Access to the operating system, which runs in the kernel, varies by the cloud model in use. Those selecting a user and kernel mode EDR may be denied access to the operating systems in some cloud models, such as the public and private HCP cloud offerings. Such nuances concerning the security controls will be discussed in the related sections of this paper. It is clear that the upper-layer applications must trust the underlying infrastructure to achieve a ZTA, so with ZT it is recommend that greater access and visibility be provided.

4. 5G CLOUD ZERO TRUST SECURITY CONTROLS

4.1 ZT Elements of 5G Cloud Deployments

4.1.1 Securing the Cloud Deployment

Regardless of the 5G Cloud Deployment Model used, the HCPs, vendors, and the relevant standards bodies (e.g., 3GPP, GSMA, IETF) must collaborate and partner to more effectively evolve from perimeter-based defenses to a ZTA. The threat models must assume that breaches will occur and the threat actor is already inside the network. Deployment models must focus on protecting the NFs, interfaces, flows, and data, supported by continuous monitoring and verifications.

Industry partnerships through standardization can help ensure that the cloud-specific threats and risks are known and understood to define effective security requirements, that vendors develop secure by design and default platforms, and that the HCPs implement layered security controls and monitoring. Currently, the industry standards bodies are struggling to find consensus to ZT-related threats, risks, and mitigations. The industry is continuing to develop comprehensive ZT security monitoring capabilities for environments with disaggregated compute stack used in virtualized compute functions (e.g., SBA/5G Core).

Implementing ZT security controls within the cloud deployment models can be quite challenging. Some security controls are going to be easier to implement (e.g., transport encryption, micro-segmentation), while others will be much more challenging to implement (e.g., 3GPP NF logging and monitoring, secure APIs, data classification). It is imperative that new deployments carefully evaluate the four deployment models referenced in Section 3.2, determine ZTA's maturity of the industry, and carefully select the deployment model that best aligns with the MNO's business, technical, and security requirements. The sections below address the ZT controls that are critical to consider when deploying 5G into one of the cloud models.

4.1.1.1 Identity Access Management (IAM)

IAM is one of the most critical ZT principles that must be carefully planned and one of the first to be implemented. Again, ZTA presumes that a threat actor is already inside the network and has gained a foothold in the network. IAM is the method to ensure that any existing footholds are removed, as well as preventing any subsequent footholds.

During the IAM planning phase, it is important to understand all the relevant required identities and accesses that are required for each NF (e.g., operating system, container, Virtual Machine (VM), application), API(s), top of rack switch/router, OSS platforms, Management and Orchestration (MANO) integrations, etc. Additionally, it is critical to identify the required and/or relevant domains and/or organizations that must be federated into IAM. The trust relationships between identity providers must be carefully evaluated before the domains and/or organizations are interconnected for identity federation.

The SBA, within the 5G Core, as already discussed, introduces enterprise-grade virtualized computing, including from cloud providers. The SBA will require advanced IAM frameworks that align with ZT principles, including dynamic access controls that are built around a “never trust, always verify” mandate. The SBA will interconnect with distributed NFs, third-party services, extend into third-party networks (e.g., private 5G, edge compute), and a plethora of UEs that include smartphones, connected autos, drones, industrial IoT, consumer IoT, etc.

For example, the following communications and interactions will need unique identities so that the NF can authorize the activities:

> Cloud Management Platform

> MANO Platform (e.g., orchestrator)

> OSS Platform

> Security Monitoring Tools (e.g., SIEM)

> API Communications (e.g., NF to NF, MANO/OSS to NF, 3rd Party to NF)

> SBI/SBA NF to NF

> Network Repository Function (NRF) to NF

> cNF to pNF

> OS to NF Process (e.g., log collector)

> NF to Other (e.g., UDM, Internetwork Packet Exchange, Mobile Virtual Network Operator, MNO)

> Human to NF (e.g., Secure Shell (SSH))

The IAM solution should incorporate continuous authentication and authorization at each layer so that a session is not established indefinitely. Least privilege access must be followed, and identity federation (e.g., Security Assertion Markup Language (SAML), OpenID Connect (OIDC), Single Sign-On (SSO)) must be incorporated so any local accounts can be eliminated or reduced.

4.1.1.2 MFA for Human Access

MFA is one of the critical features that is expected to be an integral part of an IAM system. This section provides an overview of the critical requirements of MFA in the telco cloud deployment and IAM solutions for the network management and operations functions. It also provides the various related key aspects such as the benefits of MFA, the MFA implementation options, HCP MFA offerings, MFA integration challenges, and MFA adoption in different types of cloud deployments.

The IAM service provides granular access control for telecom cloud services and related telecom network automation. IAM in cloud environments identifies (for authentication) and authorizes the cloud users to allow access to various services (related to platform, data storage, infrastructure, application, etc.) for operational and management purposes. Specifically, NIST ZTA Tenet 6 [4] highlights that all resource authentication and authorization should be dynamic and strictly enforced before access is allowed.

5G cloud deployments implementing a ZTA are expected to have Identity, Credentials, and Access Management (ICAM) and asset management systems with MFA for human access to resources. The reason is that single-factor-authentication access (e.g., password or PIN) is susceptible to credential theft, forgery, and reuse across multiple systems. User accounts in public cloud environments are generally globally accessible and more susceptible to certain types of single-factor authentication vulnerabilities. MFA relies on two or more factors for authentication to improve access security and has greater resistance to compromise. Some of the different types of MFA can be broadly classified as:

> **Knowledge:** Something a user knows (e.g., a password or a security question). This is a common type of MFA.

> **Possession:** Something a user has (e.g., a smartphone with an access code/One-Time Password (OTP), Personal Identity Verification (PIV) card, or Fast Identity Online (FIDO) security key).

> **Inherence:** Something a user is (e.g., fingerprints, iris matching, facial recognition, etc).

Phishing-resistant MFA methods, such as FIDO, and PKI-based MFA (e.g., Common Access Card/PIV card), are considered more reliable than other MFA options, such as SMS/voice-based MFA (e.g., SMS, voice message) and app-based MFA (mobile push notifications, OTP, etc.) [24]. Phishing attacks can be fully automated to operate inexpensively at scale to obtain passwords, one-time codes, and other information for illegitimate access. Thus the ICAM systems in cloud deployment should consider phishing-resistant MFA methods [25].

It is required to use MFA for cloud accounts that can be elevated to access Kubernetes clusters in the cloud because using MFA can reduce the risk in case an adversary achieved valid credentials to an account that has permissions to the Kubernetes cluster. For secure access control to the Kubernetes API, the API access should be considered privileged and should use MFA for all user access [26], [27].

One of the main challenges with MFA integration is proprietary APIs provided by partner companies rather than standardized APIs to achieve this integration. The problem with this approach is that if an API behavior is updated, integrators are required to update their products in response. This requires early notification of modifications within APIs, which may affect compatibility between products and may add additional burdens on vendors and consumers. For example, vendors need to spend resources to change their products, and consumers need to apply updates to multiple products. Additionally, MFA vendors are required to implement and maintain wrappers for each partner component to maximize compatibility and interoperability. For example, many MFA product vendors are required to create a different wrapper for each cloud provider or identity management system to be usable in different kinds of client combinations [13].

Depending on the type of telco cloud deployment, either the on-premises MFA and/or cloud MFA can be leveraged to ensure security during resource access. Some of the MFA solutions offered by HCPs includes Microsoft Entra, which works based on two or more authentication methods, such as a password, a trusted device that's not easily duplicated (like a phone or hardware key), and biometrics such as a fingerprint or face scan [28]. Human user identities include identities for employees, partners and customers, and non-human user identities include identities for applications and services, referred as workload identities.

The AWS MFA performs a second authentication factor in addition to username and password sign-in credentials. AWS MFA is enabled at the account level for root users, and the associated IAM users are then created by the account. The users can also use a federated identifier, where authentication is done based on corporate credentials and MFA configurations. Some of the MFA methods available for IAM include passkeys and security keys based on FIDO standards, virtual authenticator apps (based on Time-based OTP (TOTP) algorithm), and Hardware TOTP tokens [29].

Google follows a phased approach for MFA rollout by encouraging MFA adoption starting November 2024 and then a mandatory requirement of MFA for password-based logins by early 2025. The end of 2025 is targeted to extend MFA requirements to all users who federate authentication into Google Cloud [30].

In order to implement a ZT approach for 5G enhanced security, telco networks should take advantage of the MFA security features offered by the HCPs to secure all user access related to operations and management. In the case of hybrid environment, on-premises identification services should also integrate MFA security solutions. An on-premises MFA can also offer more control over the security operations.

4.1.1.3 PKI-Based Mutual Authentication for Machine-to Machine Communications

PKI-based mutual authentication is a well-known and widely deployed security control to provide automated, secure M2M connections in cloud environments and 5G cloud implementations. Secure protocols leverage PKI X.509 certificates issued by trusted Certificate Authorities (CAs) to authenticate both ends of the connection in a client and server manner.

Certificate Management Protocol (CMP) [31] is the standardized protocol for managing PKI certificates for 5G networks. CMP is used to handle the lifecycle of certificates required for secure communication, authentication, and encryption. It automates the primary aspects of certificate management, such as requesting, renewing, and revoking certificates. Two important specifications governing this area are TS 33.501 [32], which is the overarching security architecture for 5G, and TS 33.310 [33], which details the authentication framework for the network domain. One example is the SBA in the 5G Core, which utilizes X.509 certificate mutual authentication using mutual TLS (mTLS) over HTTP/2 with OAuth2-based authorization. Other non-SBA interfaces also use PKI based mutual authentication, such as RAN-to-core transport interfaces with IPSec and Datagram Transport Layer Security (DTLS).

For 5G M2M communications, the Root CA with relevant sub-CAs are owned by the 5G MNO. These process Certificate Signing Requests (CSRs) from the various telecom vendor network elements so the vendor solution can act as a sub-CA under the 5G MNO CA hierarchy. The infrastructure vendor then handles certificate lifecycle management for all the network elements that it manages using CMPv2. Network elements and applications that are not under a infrastructure vendor sub-CA may send their CSRs directly to the MNO CA itself.

5G NFs and applications use PKI-based mutual authentication for M2M as defined in 3GPP standards. Certificate-based mutual authentication must also be implemented across the cloud. All four of the cloud deployment models support this. With the multi-vendor stack model, the MNO is in control of the entire infrastructure and issues certificates within its existing CA infrastructure. The telco full vendor stack model often implements its solution as a sub-CA underneath the MNO CA. The HCP public and private cloud models offer advanced frameworks for PKI-based M2M mutual authentication. These models leverage cloud-native tools, managed services, and integration with external MNO PKI solutions. HCP services are HCP specific and include AWS Certificate Manager, Azure Key Vault, and Google Certificate Authority Service (CAS).

Of note is the recent addition of the Automatic Certificate Management Environment (ACME) protocol for the automated management of digital certificates in 3GPP Release 19 [34]. ACME was added to Release 19 in an informative annex as an option to avoid potential limitations of CMPv2, which included complexity associated with managing certificates across multiple types of cloud environments.

The mobile industry is closely monitoring and planning for Post Quantum Cryptography (PQC) migration using the latest relevant standards from NIST, IETF, and other standards organizations. These initiatives will mitigate the potential threat of quantum computers against asymmetric cryptography (RSA, Elliptic Curve Cryptography (ECC)), which is widely used today in security protocols such as for TLS, IPsec, SSH, CMPv2, etc.

4.1.1.4 Sensitive Data Encryption

Data-in-transit encryption using secure protocols is defined in 3GPP TS 33.501 [32] and TS 33.210 [35]. Other protocols not listed are also supported by infrastructure vendors for 5G MNO use. The security control implementation in 5G networks is well understood and utilized by 5G MNOs for the management and control plane connections. The data encryption protocols that are typically available for use include TLS, mTLS, DTLS, SSHv2, IPsec, SFTP, and FTPES. In addition to those listed above, other VPN protocols may be used in the cloud infrastructure, such as OpenVPN, SSLVPN, and Wireguard along with gRPC with TLS.

Data-at-rest encryption falls into a few different categories, including disk/volume encryption for individual servers, network storage encryption, and the Key Management System (KMS), which manages storage and use of secret keys used by the Kubernetes system. All three categories are supported in each of the four cloud deployment models to varying levels and with implementations often unique to the cloud vendor.

The multi-vendor stack and infrastructure vendor full stack model are normally based on standard Linux features that provide the data-at-rest encryption functionality. Such features include Linux Unified Key Setup (LUKS), dm-crypt for block-level encryption, fscrypt for file-level encryption, Ceph with encryption for networked storage, and integration with an external KMS. That KMS can be a common 5G MNO internally deployed solution such as Hashicorp Vault or one of the services belonging to the HCPs. The implementations are vendor specific and often must be installed at cloud infrastructure installation time. Depending upon implementation, it may be difficult or not possible to turn on and enable data-at-rest encryption after installation is complete. The HCPs for the public and private cloud offerings support unique implementations and services for server-side encryption, storage encryption, and a KMS service. Proper planning for this security control implementation is necessary.

4.1.1.5 Secure Software Development Based upon DevSecOps

Developing and executing on a Development, Security, and Operations (DevSecOps) program is critical to reducing the introduction of new cyber threats in new software releases. DevSecOps is a shift-left security strategy where security is integrated into the early stages of the Software Development Lifecycle (SDLC). DevSecOps addresses secrets management, ZT, container security, security testing, security monitoring (logs and metrics) capabilities, and security configuration monitoring embedded within the software. NIST's Secure Software Development Framework (SSDF) is captured in Special Publication 800-218 [3]. It provides a set of recommendations relating to protecting the software, producing well-secured software, and the detection, response, and remediation of software vulnerabilities.

CI/CD is a concept that has been adopted and subsequently published in NIST SP 800-218 [3], ISO 27001 [36], Open Web Application Security Project (OWASP) [37], Cloud Security Alliance [38], Software Engineering Institute [39], MITRE [40], and European Union Agency for Cybersecurity [41]. CI/CD is critical for organizations to adopt because it is intended to find security vulnerabilities and enable these vulnerabilities to be remediated prior to the software being compiled and released. For example, CI/CD can be leveraged to build these cybersecurity capabilities in the software development process:

- **Early Detection of Security Vulnerabilities** through the introduction of automated security scans such as Static Automated Security Testing (SAST), Dynamic Automated Security Testing (DAST), Dependency Scanning, Software Composition Analysis (SCA), and more.
- **Secure Secrets Management** avoids hardcoded credentials inside the source code and ensures that Role-Based Access Control (RBAC) is utilized.
- **Compliance and Governance Policy Enforcement** by validating software code complies with security policies, such as those published by NIST and OWASP [42][43], and maintains a log of all software version changes, deployments, security scan results, etc.
- **Continuous Security Monitoring and Threat Detection** by leveraging runtime security scanning and leveraging automated threat intelligence feeds.
- **Software Supply Chain Security** by requiring code signing and confirming the integrity of software artifacts. It can also be used to track the provenance of the source code, including software consumed from third parties and open-source repositories.
- **More Frequent and Efficient Patching** by using automated patch-management capabilities and automated rollback mechanisms.
- **Embedding ZT into the Software** by ensuing that each component authenticates before interworking to reduce insider threats and unauthorized access.

Following DevSecOps best practices as outlined in SSDF and building in a CI/CD pipeline will ensure that security is integrated into every phase of software development. This significantly reduces the potential of security risks being introduced into an MNO's network.

4.1.1.6 Automated Security Testing/Configuration Validation

Security vulnerability scanning and system configuration checking are both critical controls for maintaining secure 5G systems, which are primarily based on Linux, Linux Containers, and cloud infrastructure environments. These involve identifying, analyzing, and mitigating vulnerabilities in software, configurations, and system settings in order to prevent potential successful exploits.

Software vulnerability assessments look at the underlying operating systems, the software, and running applications that may be in containers or not. Any issues found are checked against a known Common Vulnerabilities and Exposures (CVE) database, such as [44]. System configurations are checked to see if they adhere to published security best practices and hardening guidelines from organizations such as the Center for Internet Security (CIS) Benchmarks [45], which are the most widely used. Automation is vital for these systems to handle the scale and complexity of modern 5G cloud environments and to be able to feed the information into other cybersecurity systems.

MNO security teams use a variety of tools, which are often a combination of commercial vendors and open source. The Security Content Automation Protocol (SCAP) is a set of NIST standards [46] to automate and standardize the identification and evaluation of vulnerability management checks, configuration checks, and for compliance reporting. SCAP is used in security scanning tools and supports automated exchange of information between different systems.

Scanning tools can focus on security vulnerability management or configuration checking or both. Vulnerability assessment tools use a privileged account to log into the system and glean the proper information, such as software versions in use and security configuration settings. Currently, best practice is to use a dedicated account for authenticated scans. The 5G MNO security team will operate a tool that is unique to their network and fits with their cloud deployment model.

Linux system security testing open-source tools include Lynis, OpenSCAP, and others. Each of the HCPs have their own cloud-specific security services: AWS Inspector [47], Azure Security Center [48], and Google Cloud Security Command Center [49].

Vulnerability assessment tools can run in most of the cloud deployment models except for public cloud offerings. The 5G MNO cannot scan any portions of the cloud infrastructure that are in the cloud provider's portion of the shared responsibility model because there is no visibility into those areas. The 5G MNO has no knowledge of vulnerabilities, secure configurations, or details of security controls deployed and utilized in HCP infrastructure, which can complicate risk management for critical infrastructure deployed in HCP premises.

One other note is that the vulnerability scanning tools have moved into a stage beyond authenticated scans where they use software agents embedded in the operating systems and Linux containers. In the context of the four cloud models, such agents are possible only in all cloud components in the multivendor stack where the 5G MNO owns all the cloud component parts. In the full vendor stack, the infrastructure vendor must agree to the agent deployment in the operating systems running the cloud services, which may require an integration effort and possibly introduce warranty and support issues. HCPs do not allow outside agents to run in the portion of the cloud infrastructure that is their responsibility, nor do they provide visibility into these layers.

4.1.2 Microservices, Trust Zones, and Micro-segmentation

4.1.2.1 Microservices

Cloud-based computing deployments allow for sets of defined 3GPP functions to be implemented using various common microservices. Microservices break down applications into smaller, loosely coupled services that can be developed, deployed, and scaled independently. Each microservice typically handles a specific business function.

Individual microservices can be deployed as multiple independently managed workloads that allow a NF to scale dynamically with network load. Cloud-based architectures can provide the necessary infrastructure and capabilities to easily deploy, scale, and manage independent microservices that can be used to implement one or more 5G NFs.

For 5G ZTA, a microservice-based cloud architecture presents challenges when implementing ZT capabilities. For example, at a high level, one might assume that each 5G NF presents a single trusted perimeter from a ZT perspective. However, a NF may be implemented using a set of common microservices operating as separate workloads. These workloads are often distributed across trust boundaries, without a single centralized controller managing the different identities or authorization capabilities. NFs may then be composed of multiple workloads that from a ZT perspective need to authenticate with one another before making authorization decisions based on the original caller, their context, and the actions of other workloads that acted on a transaction. This presents a highly complex architecture, which may be difficult to fully support within a ZTA.

A new IETF working group, Workload Identity in Multi System Environments (WIMSE), addresses the challenges associated with implementing fine-grained, least-privilege access control for workloads deployed across multiple service platforms, spanning both public and private clouds. The work will build on existing standards, open-source projects, and community practices, focusing on combining them in a coherent manner to address multi-service workload identity use cases.

Microservices architecture offers several benefits for implementing ZTA principles:

- **Security**: Each microservice can be treated as a separate security segment, allowing precise access control and limiting breach impact. Microservices ensure each service has only necessary permissions for its specific function. Mutual TLS (mTLS) provides strong authentication and authorization between services.
- **Threat Detection and Response:** The isolated nature of microservices simplifies the detection and response to security breaches. Uniform security policies can be implemented in multi-cloud and hybrid setups.
- **Operationalization:** Microservices can be updated or deployed separately, reducing risk and enabling quicker security updates. Microservices can leverage containerization for additional security and deployment ease in cloud environments.
- **Microservices architectures and trust zones** are intrinsically linked through the principle of granular security segmentation. In cloud-native environments, microservices — independently deployable services communicating via APIs — require trust zones to enforce ZT principles, isolate attack surfaces, and prevent lateral movement.

4.1.2.2 Trust Zones

A “trust zone” [4] is a set of segments within a network where resources are grouped based on similar security requirements and trust levels. NIST’s “Zero Trust Architecture” [4] emphasizes that (1) no network location is more trustworthy than another, and (2) trust should never be assumed based on physical or network location.

The evolution of ZTA has introduced critical distinctions between implicit and explicit trust zones, reflecting a paradigm shift from perimeter-based security to resource-centric protection. Implicit trust zones, rooted in legacy network designs, assume trust based on location or ownership, contradicting ZT principles. In contrast, explicit trust zones require continuous validation of identity, context, and risk posture for every access request. Access is granted based on real-time factors like user role, device health, and geolocation. Trust boundaries are drawn around individual resources (e.g., data, applications, services, systems) rather than network segments. The focus is on protecting specific resources rather than network segments or perimeters. Explicit trust zones represent a critical evolution in securing 5G NFs and cloud networks and operationalizing ZTA principles to address the dynamic, decentralized nature of modern cellular infrastructures. Explicit trust zones identify where continuous authentication and context-aware authorization could be used at the level of individual NFs. Micro-segmentation is then used to isolate resources, with software-defined boundaries replacing network-based segmentation.

The trust zone paradigm aligns with the SBA of 5G Core (5GC), where NFs are disaggregated into scalable modular components. Explicit trust zones replace broad network-level trust with per-resource, continuously validated access policies. Explicit trust zones enforce micro-segmentation at the NF level, isolating functions like the Policy Control Function (PCF) or Unified Data Management (UDM) into distinct security domains. This contrasts with legacy 3G/4G architectures that treated entire network planes (e.g., EPC) as trusted. In addition, identity-based segmentation creates software-defined perimeters, allowing NF-to-NF communication only after mutual authentication via TLS 1.3 with certificate pinning. SBA Domain Security is provided by 3GPP TS 33.501 [32], which mandates OAuth 2.0 for NF service authorization, ensuring each API call between NFs validates the consumer’s credentials and scope.

The use of Linux containers can significantly enhance the implementation of ZTA principles. A microservice or microservices are built within a container, which is a lightweight runtime environment. Containerization supports fine-grained micro-segmentation and provides the granularity, flexibility, and consistency needed to effectively implement the core principles of ZTA as defined in [3]. Containerization can be tightly integrated with other ZTA components, such as IAM, Policy Enforcement, and Continuous Monitoring, to enhance overall security.

4.1.2.3 Micro-Segmentation

Micro-segmentation in a cloud environment is best understood by breaking it down into three main areas:

container orchestration, Container Network Interface (CNI), and service mesh.

A container orchestration platform provides automation and management for the Linux containers running the 5G NFs and applications. Kubernetes is the dominant container orchestration platform.

CNI provides both the networking interface and segmentation features that enforce network policy at OSI networking layers 3 and 4 for the container. A CNI uses Linux namespaces to isolate containers and pods within the larger Linux host server. Calico and Cilium are the primary CNIs used today. Calico uses IP tables and extended Berkeley Packet Filters (eBPF) as the underlying Linux technology to enforce policy. Cilium uses eBPF. Each can implement standard Kubernetes network policies for micro-segmentation. Each has additional unique policy options.

Service mesh is a dedicated layer that manages communications between containers and provides advanced ZT security controls at OSI network layer seven, such as mTLS for secure authenticated and encrypted communications between containers. A service mesh is made up of a data plane with sidecar proxies, which intercept the traffic and provide these advanced security mechanisms, along with a control plane for security policy management. Popular service meshes include Istio and Linkerd.

In the multi-vendor stack model, the 5G MNO will select the container orchestration platform, CNIs, and the use of a service mesh. With the infrastructure vendor full stack, the infrastructure vendor selects and implements all three. The HCPs' public and private cloud offerings have their own services that implement all three in their CaaS offerings. For example, Google [50] and AWS [51] offer Calico and Cilium CNIs, whereas Azure [52] offers Calico and the Azure CNI. Each has its own service mesh offering. 5G Cloud software architects must take care to ensure that any micro-segmentation of microservices are deployed consistent with the ZTA principles across the entire system.

In summary, micro-segmentation can be used to create fine-grained, isolated security zones within a network. For the 5GC, this means treating each NF — and even components within an NF — as a distinct segment with its own specific security policies. By defining strict rules for communication between segments, micro-segmentation enforces the principle of least privilege. Each NF is allowed to communicate only with the specific NFs and services it absolutely needs to function, thus minimizing its potential attack surface. Micro-segmentation provides better visibility into the communication patterns within the 5GC, allowing for more effective monitoring and control of network traffic.

If one NF is compromised, micro-segmentation significantly limits the attacker's ability to move laterally to other parts of the 5GC, which is important to defend against evolving Advanced Persistent Threats (APTs). The strict communication boundaries prevent the attacker from easily accessing other NFs or sensitive data.

4.1.3 Information Input Sources to Enable ZT Policy Decisions

To supplement the security controls described in the section above, this section will address the various capabilities needed as information sources to enable ZT policy decisions.

4.1.3.1 Continuous Monitoring, Logging, and Alerting

Telecom infrastructure has diversified when it comes to 5G with respect to the architectural assets and use cases. Various standards have modularized the entire system, bringing in greater flexibility, but at the same time increasing the potential attack surface area. Security plays a bigger role in ensuring that all aspects of the infrastructure uphold the triad of confidentiality, integrity, and availability to the maximum possible levels.

Traditional software security has typically been static, with the assumption that the threats are mitigated once addressed. But today's adversaries are continuously finding new ways to compromise the security of the infrastructure at the slightest possible drop of guard. Hence newer mechanisms need to be explored where threat detection can help to stay one step ahead of such adversaries. This introduces the need for Continuous Security Monitoring (CSM), which assumes that the system is always at risk and hence keeps probing continuously to detect signs of security degradation. Based on information gathered by continuous security monitoring, further steps can be undertaken to assess for any potential threats and, when identified, execution of appropriate mitigations.

The ZT concept assumes that security is never taken for granted, and thus all users, assets and resources are continuously validated. The NIST SP 800-207 Zero Trust Architecture publication [4], which defines the seven tenets of a ZT architecture, specifically calls out the need for CSM in tenets 5 and 7 which state:

> **NIST Tenet #5:** The enterprise monitors and measures the integrity and security posture of all owned and associated assets.

> **NIST Tenet #7:** The enterprise collects as much information as possible about the current state of assets, network infrastructure, and communications and uses it to improve its security posture.

The challenges to achieve CSM should be investigated further, keeping in mind both MNO and infrastructure vendor perspectives. SDOs like 3GPP and O-RAN have been studying this topic while focusing initially on the SBA Core and RAN, respectively. Both are looking at further expanding their studies into other domains they are responsible for with a goal of reaching a consensus on what needs to be standardized.

The question of what exactly needs to be standardized, and what is left as an implementation decision, is spread across many aspects of CSM, as described below. The basis of CSM is the data that needs to be identified, defined, collected, exposed, and assessed to detect threats in a system. Then the system's security posture must be understood in terms of whether there is a quantifiable baseline and how dynamic policies can be invoked in relation to the mitigation of the threats that identified. The following discusses each of these aspects in more detail.

- **Definition:** The most impactful threats are listed and the corresponding data for each, as a minimum, is identified. This data becomes the basis for the evaluation of a threat. Consideration is done as to whether the data definition needs to be standardized and to what extent. Data could, for example, be statistical data like counters and event-related data that must be defined so that it can be consumed.
- **Collection:** Having defined the data for CSM, the next step is to check if a collection mechanism is required within the system that needs to be standardized. It also must be decided if a correlation mechanism is required to group data logically related to a particular threat or family of threats.
- **Exposure:** After collection, the data has to be sent to the external systems for consumption. Whether existing standardized interfaces and frameworks suffice, along with the nature of the streaming formats, are then explored while at the same time keeping an eye on the performance impacts of sending security data in addition to existing non-security data.
- **Evaluation:** Once the data has been collected by a system, an evaluation needs to be carried out to normalize varying standards of data structure and formats for specialized systems like SIEMs and SOAR.
- **Monitoring:** Data must be continuously ingested and retained for various time periods in accordance with 5G MNO policy.
- **Mitigation:** Remedial actions on the victimized endpoints or NFs based on the threats are executed.
- **Security Posture:** The MNO defines their overall and complete security posture, which then becomes a baseline to check and measure threats.
- **Dynamic Policies for CSM:** A policy must be defined. For example, one policy could be for the initial configuration of data points needed to collect data are defined. Then, any mitigation actions are informed to the impacted assets.

In Figure 5, the left side shows the challenge the industry faces where currently security data is of an unstructured format and not standardized. The right side shows a future scenario where security data is structured and/or standardized (e.g., to some degree for most important threats). Having structured and/or standardized security data would be the ideal outcome. However, the reality is that industry will be challenged to adopt and achieve this goal in the foreseeable future due to the considerable impact it can have on deployed, and soon to be deployed, systems.

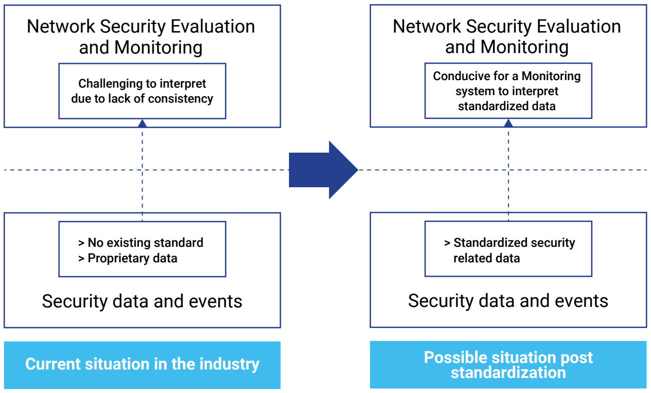

Network Security Evaluation and Monitoring
Challenging to interpret due to lack of consistency
> No existing standard
> Proprietary data
Security data and events
Current situation in the industry
Network Security Evaluation and Monitoring
Conducive for a Monitoring system to interpret standardized data
> Standardized security related data
Security data and events
Possible situation post standardization

Figure 5: Continuous Security Monitoring Development Opportunities

Structured data would be required from a micro and a macro view for visibility across multi-vendor environments. This also leads to different possible deployment options, keeping in mind that vendors would be able to standardize some aspects of CSM, while at the same time MNOs would like to leverage their existing, and sometime proprietary, CSM capable assets.

Generating structured and/or standardized security data is one part of the challenge. Other aspects must be considered, such as the collection and exposure of security data to external systems including SIEM/SOAR. Other topics that need consideration include having each 5G network entity generate and expose structured and/or standardized security data externally, or having an intermediate network entity that collects security data from multiple internal entities and exposes it externally as structured and/or standardized security data. The current state of security data from the cloud perspective is discussed in the next section. The 3GPP viewpoint and work on this topic was discussed in Section 2.1 on the current state of ZT in 3GPP.

4.1.3.2 SIEM/SOAR

The Security Information Event Management (SIEM) and the Security Orchestration Automation & Response (SOAR) are the two central monitoring, detection, and automated response systems of the 5G MNO's security operations center. The SIEM collects, processes, and correlates security event data such as logs and telemetry to detect and alert for security threats. The SOAR automates and orchestrates the processes used by incident response personnel to promote the fastest possible response to intrusions. Both the SIEM and SOAR operate in close cooperation and represent a large monetary investment and require highly skilled cybersecurity personnel. Due to the large amounts of investment required, the SIEM/SOAR are centralized. They must cover multiple cloud deployment models and handle the security detection and response to threats for the 3GPP 5G NFs/network equipment in addition to the deployed cloud environments.

These systems developed in the IT market space and as a result lack 5G cellular-specific feature functionality out of the box. Two responses developed from this situation:

- Infrastructure vendors developed products that met the use cases that were unique to the telco environment, such as threat management for the

RAN and core NFs. These solutions detect threats and then report security-related notifications to the primary SIEM/SOAR and thus operate as a cellular adjunct or branch.

> If such solutions are not deployed, then the second response is needed. Here, experienced cybersecurity personnel can write custom rule sets unique to the infrastructure vendor product logs and telemetry for the SIEM. This is a complex task that requires expertise in infrastructure vendor equipment logs, log formats, and telemetry sources, along with a deep understanding of RAN and core functionality. Only the larger cellular service providers can afford the investment and personnel to customize the SIEM to their environment using the above two solutions at present.

Turning to the cloud deployment models, the logs and telemetry from the various portions of the cloud technology stacks tend to be supported by the SIEM/SOAR vendors and are much more easily consumable because these are central to the IT world.

In the multi-vendor and infrastructure vendor full-stack solution, the 5G MNO owns the cloud infrastructure, so all required logs and telemetry from the hardware, network, and host operating systems can be sent to the SIEM. With the public and private cloud offerings, the logs and telemetry offered by the cloud provider's logging services must be sent to the SIEM. Each HCP has a different logging offering. The SIEM must have the feature set to process them, but this ability is common due to the ubiquity in the use of cloud services. Full visibility into the underlaying cloud infrastructure is not possible in the layers of the cloud stack that are the cloud provider's responsibility.

eBPF is as a telemetry source in the Linux kernel of the host Linux operating systems that can assist SIEM/SOAR with visibility in 5G cloud deployment models. eBPF runs in the kernel and sends many types of telemetry to the Linux user space without a large impact on performance including logs, system calls, process information, network packet information, CPU and memory usage, security events and anomalies, and others. Such a stream gives a full view into the various containers running the 5G NFs and can be tuned to the data of interest.

eBPF gives a complete and detailed view of all activity without the performance impact of legacy methods such as Deep Packet Inspection. eBPF can be utilized in cloud models where the MNO has access to the Linux host operating systems, such as the multi-vendor model, and IaaS for the public and private cloud. Public and private cloud PaaS offerings do not allow such access. With the infrastructure vendor full-stack solution, the infrastructure vendor would be required to allow eBPF installation and use, which likely would require testing. Rulesets and ML/AI for eBPF for Linux are built by the SIEM/SOAR vendors and can be tuned by the 5G MNO cybersecurity personnel.

The SIEM/SOAR vendors currently support threat intelligence feeds based around IT security frameworks such as MITRE ATT&CK, which has extensive industry support. Two frameworks specific to the telco world have been developed and continue to evolve: MITRE FiGHT and GSMA MOTIF. SIEM/SOAR vendors are encouraged to support these two frameworks.

4.1.3.3 Anomalous Behavior Detection using AI/ML

The Enduring Security Framework (ESF) recommends anomalous behavior detection for multiple roles in detecting attacks, preventing lateral movement, and isolating network resources within the 5G cloud infrastructure. The ESF further suggests that the massive amount of network traffic and identity and access management events in a 5G cloud infrastructure requires AI/ML to detect anomalous adversarial behavior, including malicious use of networks, accounts, and other customer cloud resources [53],[54].

More specifically, AI systems can analyze user behavior, device interactions, and network traffic in real time, enabling continuous monitoring that is crucial for ZTA. AI can identify anomalous patterns that may indicate security breaches or unauthorized access attempts. This capability aligns with the ZT principle of "never trust, always verify," because AI provides ongoing assessments of user and device trustworthiness based on current context rather than static credentials [53].

ZT dynamic policy enforcement for isolating network resources calls for real-time threat detection of how well pod/container isolation is working and real-time response if the isolation is breached. Anomalous behavior detection can play a role in real-time threat detection and response by looking for variances from the defined allowed behaviors of authenticated and authorized users and systems [54].

A 5G Americas white paper offers insights into AI/ML benefits for 5G networks while considering the potential risks in AI/ML opening new attack vectors on those 5G networks. New attack vectors on 5G networks can come in two varieties: attack vectors directed at the AI/ML system and AI/ML-enabled attacks on a 5G network. Countering these attacks starts with employing currently available best security practices followed by emerging security control frameworks addressing the known AI/ML threat landscape. AI/ML systems need security governance comprising policies, guidelines, and compliance with enterprise and regulatory requirements. The white paper describes several possible AI/ML-enabled sophisticated attacks on mobile networks and cites the unknown feasibility of these approaches [55].

AI/ML excels at recognizing complex patterns and subtle anomalies that traditional monitoring systems might miss. The steps in the automated anomaly detection process are anomaly identification, validation, impact assessment, alerting, data collection for accuracy measurement, and AI/ML model update. The updates refine the system's understanding of normal behavior based on new data and feedback to ensure the model evolves in response to changing patterns and new types of anomalies.

AI-based threat hunting for cloud operations is an advanced cybersecurity approach that leverages AI and ML algorithms to *proactively* identify, analyze, and mitigate potential security threats in cloud environments. Unlike traditional threat detection systems that rely on known attack patterns, AI-powered solutions can identify previously unknown or emerging threats, such as zero-day vulnerabilities. This typically relies on identifying patterns and anomalies that might indicate potential security breaches in conjunction with behavioral analysis. AI systems learn the normal behavior of an organization's network, applications, and users, establishing a baseline that enables alerts to be raised in real time when deviations from this baseline are observed. This also enables prediction of future vulnerabilities by analyzing how programs interact with files, networks, and system resources. It can monitor network traffic for irregular patterns suggesting data exfiltration.

The SIEM is one of the 12 fundamental security control groups and central to the four cloud models mentioned in this paper. The SIEM could also play a central role for AI/ML processing of the massive volume of monitored security data and detecting anomalies in the security data. Further AI roles in the SIEM could prioritize alerts based on risk, translate human queries into their processing language, and assist in various points in the incident response workflow processes.

Recommendations for MNOs and cloud service providers:

> Consider AI-based anomaly detection for enhanced visibility of malicious activity in 5G infrastructure.

> Future work should be initiated to explore how to defend against AI-based attacks on 5G cloud infrastructure.

4.1.3.4 Threat and Endpoint Detection and Response (TDR/EDR)

Telco critical infrastructure nodes or endpoints are engineered to run a telco-specific workload, such as a 5GS Access and Mobility Management Function (AMF) or cloud RAN Centralized Unity (CU) or Distributed Unit (DU). The telco node environment has a static nature in that software configurations and applications are not modified daily. The job of a telco EDR solution is very specific to protect a telco workload running in a Linux container or virtual machine. Telco EDR solutions are engineered to know a telco workload's expected behavior and monitor its actual behavior and its environment. The EDR can then detect and alert of any potential threats/anomalies to enable mitigation actions in response.

A telco EDR solution typically consists of a centralized telco EDR server communicating with and managing telco EDR agents deployed alongside telco workloads on multiple endpoints. The telco EDR server can be considered the central collection point of the solution, while the distributed telco EDR agents feed the EDR server. In the IT world, this correlates to EDR agents running on server applications in a data center or in the cloud.

Telco EDR agents run on the same endpoint alongside the telco workload that they are tasked with protecting. Thus the performance impact on the telco workload endpoint is minimized. The telco EDR is monitoring an endpoint, which is single purposed and is modified only during system/application upgrades. As a result, the telco EDR can focus on endpoint application and configuration consistency. By keeping the telco EDR agent design restricted to a focused purpose, frequent updates are not required because most of the detection intelligence is implemented on the centralized telco EDR server, where the detection rules are pushed to the EDR agents. Telco nodes are critical infrastructure, so enabling automatic mitigations within the EDR agent should be carefully evaluated by the MNO before enabling. This approach reduces the possibility of a false positive triggering an action that could affect the node's functionality and performance.

EDR agents are not immune to security threats and must be designed, deployed, and operated with stringent security controls in place. It should also be acknowledged that EDR solutions in the industry adopt different approaches for allowing their EDR agent to run in Kernel mode or user mode. Both have their own pros and cons, which are not elaborated on in this paper. When choosing a solution, a key consideration is security, particularly the ZT principle of least privilege and system stability.

In a November 12, 2024, report, "2023 Top Routinely Exploited Vulnerabilities" [56], CISA noted that 10 of the 15 most frequently exploited vulnerabilities were initially zero day exploits. EDR agents are a key critical security control in a modern 5G network to defend against novel attacks by advanced adversaries. EDR deployment in virtual machines is well understood.

5G networks deploy cloud-native network functions (CNFs) in Linux containers. EDR agents for 5G CNFs require rulesets that are particular to the Linux distribution, Linux security-related functionality used, and the CNF itself, and must be able to handle standardized 5G protocols. Given these requirements, out-of-the box rulesets for commercial EDR agents are inadequate and must be configured and tuned by knowledgeable personnel. Two options are available:

> Use a commercial EDR solution that has been extensively tuned to the infrastructure vendor NFs and supports the Linux environment used by the cloud. Either the 5G MNO or the telco NF vendor is responsible for the rulesets, but this option may result in telco NF vendor warranty terms being violated. Specific support for the EDR commercial vendor by the telco NF vendor may be required.

> Use a recommended EDR solution provided by the telco NF vendor for which that vendor is responsible for the ruleset configuration. The telco NF vendor is responsible, so there are no warranty issues.

There is a problem in the deployment of EDR agents that is particular to telco NFs in the cloud-native container environments. These agents run as cloud native and are required to run on Linux systems of different distributions inside Kubernetes-managed containers where the NFs are executing. This is a different agent-execution environment than the legacy virtual machines. With containers, these agents require visibility into Linux system calls, which are requests by an application to the operating system to do a task such as open a file or to access a hardware device.

EDR agent implementation varies, with some running in only user mode and others in both user and Kernel mode. NFs run in containers in user mode separate from the underlying core operating system functions that run in the Linux kernel. Kernel mode access provides broader security coverage through access to telemetry sources such as system calls, system logs, and network traffic, giving Kernel mode EDR agents a different scope of coverage than user mode EDR agents. Access to kernel sources is commonly provided by eBPF.

If the 5G MNO chooses a user- and/or Kernel-mode EDR, the architecture requires sensor code running in a Linux user space covering all the containers, or within each individual container, and a second part running as a module in the Linux kernel. As Figure 6 shows, this poses no issues when the cloud infrastructure is owned and operated by the 5G MNO in the legacy multi-vendor-stack and telco full-vendor-stack deployment models. EDR agents can be deployed normally by the security and infrastructure teams. If the cloud deployment is done with IaaS — where the 5G MNO can deploy their own Linux servers, Kubernetes stack, and individual telco CNFs — then the MNO should be able to deploy and operate the EDR agent and sensor code. If the cloud deployment is done with PaaS via a HCP using either the public, hybrid, or private models, the MNO will be challenged to get the EDR agent and sensor code deployed.

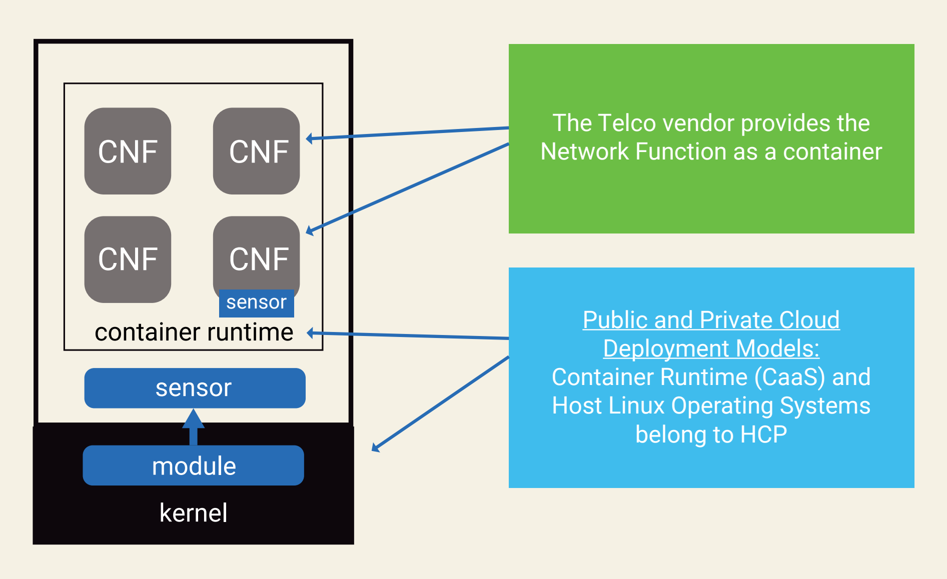

CNF
CNF
CNF
CNF
sensor
container runtime
sensor
module
kernel
The Telco vendor provides the Network Function as a container
Public and Private Cloud Deployment Models:
Container Runtime (CaaS) and Host Linux Operating Systems belong to HCP

Figure 6: Container as a Service (CaaS)

There are additional considerations when running in the hybrid HCP public cloud and HCP private offering deployment models using PaaS. In the hybrid deployment model, the HCP is providing the infrastructure, including the CaaS portion that runs the 5G CNFs. EDR agents that operate in both user and kernel space provide broader coverage of all the system calls and internal information as seen by the kernel. The user space component would be required to run in the HCP's CaaS. The user space component could also run in the NF itself and would then be the responsibility of the telco NF vendor and the 5G MNO. The kernel component would run in the Linux host operating system, which belongs to the HCP.

This approach needs clarification about shared responsibility when the 5G MNO wants to run an EDR that also operates in the kernel space. One possibility is for the HCP to provide EDR as a service to the 5G MNO within the private cloud offering stack for critical infrastructure customers. Provision must be made for the 5G MNO security operations staff to be able to configure and tune the EDR rulesets to the particular NF. The EDR agent also must be tested and supported within the HCP CaaS environment. EDR notifications and alerts would go to the EDR management system owned by the HCP and then forwarded to the 5G MNO SIEM.

4.1.3.5 Threat Intelligence

Threat Intelligence (TI) feeds provide real-time information concerning potential threats including observed Indicators of Compromise (IOCs) from actual observed attacks such as IP addresses, file hashes, and URLs used by specific threat actors. TI feeds combine with a security framework, which define real-world attacker Tactics, Techniques, and Procedures (TTPs), plus mitigations and recommended detections. Both are combined and are used by tools such as the SIEM and EDR for threat detection and visibility, threat hunting, and incident response by the organization's Security Operations Center (SOC).

In the enterprise space, a mature ecosystem exists, with hundreds of commercial and open-source TI vendor feeds that utilize the MITRE ATT&CK security framework. MITRE ATT&CK includes a cloud matrix that covers both public cloud and hybrid/private cloud architectures. In the security framework ecosystem, two standards are important to allow the automated sharing of TI information: Structured Threat Information eXpression (STIX), which uses JSON to encode information [57], and Trusted Automated eXchange of Indicator Information (TAXII), which transports STIX formatted data using HTTPS [58]. MITRE ATT&CK is supported by the IT cybersecurity ecosystem, which uses STIX and TAXII to share attack and threat information in an automated machine-readable structured manner [59].

In contrast, the 5G cellular provider ecosystem has a much smaller subset of IT vendors and feeds due to the unique nature of 5G systems that employ unique cellular protocols and have use cases not found in the IT world. Security frameworks are built on actual real-world threat and attack data, but not many 5G-specific attacks have been reported, thus making such data sparse. MITRE ATT&CK does not fully cover 5G, so there are two developing 5G-focused security frameworks based on MITRE ATT&CK: MITRE FiGHT and GMSA MOTIF. The 5G scope in both frameworks includes fraud, roaming signaling, false base stations, SMS and voice SPAM, and other unique mobile use cases. MITRE FiGHT focuses exclusively on 5G SA and NSA networks and applications. Version 1 was released in September 2022, with the current version 2.1, as of December 2024 [60]. V3.0 adds cloud-centric threats. FiGHT support is now included in some vendor products.

GSMA MOTIF version 1 was published in March 2024 [61] and covers all cellular generations, from 2G through 5G. Two documents are provided: FS.57 MOTIF Principles, which is public, and FS.58 MOTIF Examples, which covers three attack cases and is available for GSMA members only. STIX support is covered in FS.57 Annex A. GSMA has access to more non-public threat information from its members via the GSMA panel of experts and various working groups.

There is currently very good cooperation between the FiGHT and MOTIF teams as both sides contribute to the other. Both FiGHT and MOTIF are still works in progress. Additional work on both, along with support by the telecom cybersecurity ecosystem, is needed in order to reach the level of maturity that MITRE ATT&CK has attained.

4.2 Cloud-Native Policy Management for Zero Trust

Policy management is a key component of a ZTA. As 5G deployments adopt ZTA, the challenge shifts from conceptual acceptance to policy operationalization across heterogeneous multi-cloud platforms. 5G multi-cloud deployments include a highly structured 5G Core and RAN framework with standardized functional elements and interfaces plus an orchestration system to manage the operational system as it evolves under changing context and situations. These structures can be dynamic, containerized environments, each governed by distinct security paradigms that are implemented using policies.

The ZTA policy framework consist of a Policy Decision Point (PDP) and a Policy Enforcement Point (PEP) logical entities, where the PDP is made up of a Policy Engine and a Policy Administrator.

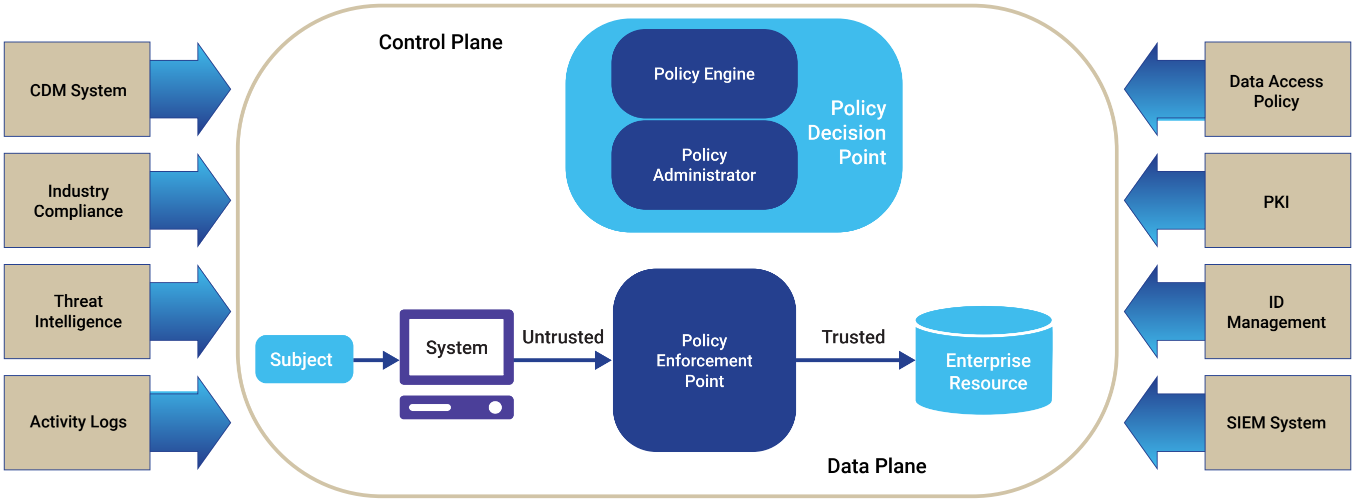


Figure 7: Core Zero Trust Logical Components from NIST 800-207 [4]

We interpret the above figure as a logical architecture that may be implemented in a variety of ways for cloud environments.

4.2.1 Policy Implementation and Deployment Strategies

There are four types of policy management:

> A **centralized** approach manages policies by a single, central authority. It enhances visibility and control because all actions are seen through a central console.

> A **federated** approach distributes management across multiple autonomous entities that collaborate while maintaining some level of independence.

> A **hierarchical** approach organizes policies in a tree-like structure, with global policies at the top and more specific policies at lower levels, with lower-level policies never violating higher-level ones.

> A **hybrid** approach combines elements of different models to balance the needs of different parts of the organization.

This report recommends a hybrid approach to ZT policy management due to the wide diversity of security needs and requirements necessary to deal with a dynamic operations/orchestration environment, as well as the highly structured and standardized 5G Core and RAN. This document will address policy management guidance for the 5G Core and RAN. However, the operations and orchestration systems do not have standardized architectures and thus introduce additional complexity. A more in-depth discussion of the policy management architecture across the entire 5G cloud system will be covered in a subsequent ATIS report targeting the topic of policy management**.**

4.2.1.1 Policy Management Guidance for 5G

In the 5G Core and RAN, 3GPP defines a set of modular network elements providing well-defined functions (as defined in 3GPP TS 23.501)[62]. The PCF is equivalent to the PDP for subscriber policies. The UDM/UDR is the PDP for subscribers. The AMF, Network Exposure Function (NEF), Security Anchor Function (SEAF), Session Management Function (SMF), and the User Plane Function (UPF) can each act as a PEP for subscriber access. All NFs are PEPs for communication internally in 3GPP networks. The AMF, Authentication Server Function (AUSF), NEF, Network Data Analytics Function (NWDAF), NRF, and UDM all contribute to the trust engine.

The Application Function (AF) interacts with the 5GC (often through the NEF) to influence traffic routing, Quality of Service (QoS), and policy. For a comprehensive ZTA, the policies enforced might be driven by the requirements and trust level of the specific applications. The AF's role in providing context and potentially triggering policy updates is important in a ZTA. The AF interacts with the PCF via N5 interface and can have the NEF in between them.

The NRF facilitates the secure discovery of NFs; this ensures that only authorized and trusted NFs can discover and communicate with each other. The NRF can play a role in this by incorporating authorization checks during the discovery process. The NRF and a CA (not 3GPP standardized) are the PDPs for authorization and authentication for internal 3GPP communication.

The Network Slice Selection Function (NSSF) selects the appropriate network slice for a UE based on its service requirements and subscription. This enables creation of isolated network environments with specific security policies tailored to the slice, aligning with the principle of limiting the blast radius of a potential breach.

The Service Communication Proxy (SCP) ensures secure communication between NFs, which is fundamental to a ZT environment within the core. As such, it can function as a PEP for facilitating secure communication between different NFs within the 5GC.

The UDM/UDR is the main PDP for subscribers, such as allowing access, which APN/DNN to use, QoS etc. The Unified Data Repository (UDR) serves as a key data source for the Trust Engine and can also store policy information.

Finally, legacy EPC NFs (MME, SGW, PGW) could also be involved for ensuring consistent security in interworking scenarios.

ZT in the 5G Core requires clear separation of these NFs. Advantageously, Kubernetes network policies can be used to restrict NFs to designated service meshes, preventing lateral movement while isolating the NF.

To support a ZTA, each NF should authenticate peers to prevent impersonation or man-in-the-middle attacks. All NFs should present certificates issued by a trusted PKI. For example, the NRF validates NF identities during service discovery and token issuance.

NRF can also be used as the PDP for ZT policies, in conjunction with the PCF, UDM, monitoring systems, and AI-driven user/device/network analytics, ensuring access decisions adapt to contextual factors.

The SCP is a NF that facilitates communication between NF consumers and producers, acting as a proxy to offload responsibilities like discovery, routing, and load balancing. Thus, the SCP can also be used as a PEP in ZTA to mediate all inter-NF API calls, blocking unauthorized requests. This enables continuous session validation.

The NEF can act as a Zero Trust Network Access (ZTNA) gateway for Application Functions (AF) and external/third-party APIs attempting to communicate with the 5G Core.

In the O-RAN Alliance architecture, the SMO is designed to support ZTA PDP functionality that can be implemented as a SMO service or as security rApps.

4.2.1.2 PaC: Policy as Code

Policy-as-Code (PaC) defines, manages, and enforces policies through code rather than manual processes. Policies are written in machine-readable formats (e.g., YAML and JSON), stored in version-controlled repositories, and integrated into CI/CD pipelines (see Section 4.2.1.4) to enable automation, consistency, and repeatability.

PaC supports automated testing and validation of policies to ensure they are error-free and conflict-free before deployment. It also enables dynamic enforcement of policies based on real-time context (e.g., user roles or device posture), aligning with ZT principles.

PaC also integrates with observability tools to monitor compliance in real time and trigger alerts or remediation when violations occur. Some of the more common PaC tools in use today include Open Policy Agent (OPA), Hashicorp Sentinel, Kyverno, and OPA-Gatekeeper.

4.2.1.3 IaC: Infrastructure as Code

Infrastructure-as-Code (IaC) is a method for managing IT infrastructure using machine-readable definition files that treat infrastructure configurations like software code. IaC supports both declarative approaches for defining desired states and imperative approaches for specifying step-by-step instructions.

IaC integrates with security tools to enforce secure configurations by default (e.g., encryption or access controls), adapts to varying load demands through autoscaling mechanisms without manual intervention, and simplifies disaster recovery by enabling rapid re-creation of infrastructure using predefined templates. These capabilities are very useful tools when deploying a ZTA. Some of the more common IaC tools used today include Terraform, which is a cloud-agnostic IaC used across Azure, Google, AWS, Ansible, Chef, and AWS CloudFormation used in AWS.

4.2.1.4 CI/CD Pipelines as the Enforcement Backbone

CI/CD pipelines are central to implementing a ZTA in cloud-based environments because they integrate security, compliance, and operational controls directly into the software development lifecycle.

For the vendors and/or cloud providers, a CI/CD pipeline can enforce strict identity verification for users, services, and automation tools. Techniques such as MFA, ABAC, and federated identity mechanisms (e.g., OIDC) ensure that only authorized entities can interact with the pipeline. Least-privilege principles minimize access to sensitive resources, reducing the risk of unauthorized actions or insider threats.

For MNOs, CI/CD pipelines are a bit more complicated in regard to the telco infrastructure that powers the mobile networks. For example, the infrastructure vendor may produce new software using their CI/CD pipeline that includes automated security testing and security controls within the software itself, but the infrastructure vendor may not have full access to the MNO's network to fully deploy the new software. In this case, the infrastructure vendor will deliver the software to the MNO, which then tests the software in a lab before rolling it out following a structured process.

4.2.1.5 Orchestration and Context

Orchestration in ZTA relies on comprehensive, real-time knowledge of all assets. This enables the orchestration system to dynamically adjust access controls and component configurations based on the current state and context of assets. Without orchestration, ensuring consistent policy enforcement and communication between these components becomes challenging to manage. An orchestrator can be a PDP that manages access decisions, whereas an orchestrator manages application or system operations. A PDP and orchestrator can have overlap, but they are not the same. The following is a list of required orchestrators in a ZTA for 5G cloud networks:

> A **High-Level Orchestrator** can be used to enable unified control over all policies, ensuring that security and compliance measures are consistently enforced. It provides a single pane of glass for monitoring policy effectiveness and compliance. It simplifies dynamically adjusting policies based on real-time data and threat intelligence.

> A **Hybrid Policy Management Orchestrator** can be used for hybrid policy management. A dedicated orchestrator is required because hybrid environments often involve multiple platforms (cloud and on-premises) that require tailored policies. It also ensures that security policies are uniformly applied across diverse infrastructure components, reducing the risk of configuration drift or inconsistencies.

> A **Microservices Orchestrator** for containers that include micro-segmentation at the CNI level.

> A **Service Mesh Orchestrator** manages, configures, and automates the deployment and operation of a service mesh, enabling secure and reliable communication between microservices in a cloud environment.

> A **Pipeline Orchestrator** can be used for governing CI/CD and similar pipelines.

> An **IaC Orchestrator** can be used for managing workloads and resources across hybrid and multi-cloud environments.

> A **PaC Orchestrator** can be used to automate the enforcement of policies across different stages of the software development lifecycle, ensuring consistent application of security and compliance rules.

> **IAM** is a human access control. This would include access to the MANO planes for the 5G NFs and the cloud infrastructure components. The PDP evaluates various pieces of context information such as the user identity, MFA status, risk score, and others. It includes a variety of access control mechanisms that can be chosen to optimize access to resources in the NF and the cloud.

> **Device Trust Evaluation** is the device that the human user is using for compliant access. This could include checks by the Mobile Device Management (MDM) and EDR agent.

- **API Access Management** covers the NEF and access to cloud APIs.
- **Data Access Control** for data or file level.
- **Policy-Based Routing or Conditional Forwarding** enables the dynamic and policy-driven control of network traffic. Orchestration provides the intelligence and automation to configure and manage these routing mechanisms based on the principles of ZT, ensuring that access is granted only to authorized entities, under specific conditions, and with the least necessary privileges.

Context awareness and situational awareness for the various policy decisions made by the PDP must be specified and obtained. The focus is on understanding the system's current state and enabling fine-grained access decisions based on immediate factors available at that point in time. Different PDP entities may use context [55] to determine the set of applicable policies that affect their behavior. In a cloud environment, resources and users are often distributed and constantly changing.

Context awareness provides more informed and dynamic access decisions based on factors like user role, device status, location, and time of access, while enabling the creation and enforcement of more precise policies and adapting to the fluid nature of cloud resources. 5G networks are inherently distributed, so context awareness helps secure the various 5G cloud infrastructure components by considering the context of each access request. Anomalies in access patterns or user behavior can be more readily identified by analyzing contextual information, enhancing threat detection capabilities in the cloud. This enhancement aligns well with the ZTA principle of continuous monitoring and verification. Each of the 12 listed above requires context information, which must be defined and then pulled from various points in the overall architecture.

Situation awareness provides a broader view of the system's current state with respect to possible future states. Situation awareness [55] is the perception of data and behavior that pertains to the relevant circumstances and/or conditions of a system or process, the comprehension of the meaning and significance of these data and behaviors, and how processes, actions, and new situations inferred from these data and processes are likely to evolve in the near future. The ZTA does *not* define situational awareness. The benefits of adding situation awareness include:

- Enhanced proactive security (e.g., by predicting how situations may evolve, organizations can take preemptive actions)
- Improved decision-making (e.g., administrators can make more informed decisions about resource allocation and security policy adjustments in the cloud)
- Adaptive risk assessment (e.g., by adjusting security measures based on the current state of the cloud environment)

Situation awareness enables predictive and proactive security measures to ensure system goals are met in the face of changing context. This enhancement extends ZTA's capabilities beyond its current scope to provide a more predictive security posture.

The PDP should be defined for all applicable parts of a ZTA solution in a 5G cloud environment. The definition of the required context and situational awareness for those PDPs is complex and on-going. ATIS recommends follow-on work to define the policy architecture for 5G ZTA.

5. CONCLUSIONS, RECOMMENDATIONS, AND NEXT STEPS

The implementation of ZTA within 5G cloud environments presents both opportunities and challenges. As 5G networks transition to cloud-native deployments, the ZT principles become essential to securing critical infrastructure, ensuring data integrity, and mitigating emerging threats. This document highlighted the various 5G cloud deployment models, the security controls necessary for ZTA adoption, and the challenges associated with securing both cloud-native and traditional architectures.

The complexity and the hybrid nature of 5G cloud infrastructures necessitate a multi-layered security approach that integrates continuous monitoring, identity management, micro-segmentation, threat intelligence, and automated security controls. Although industry efforts to align 5G security with ZT principles are underway, significant gaps remain in standardization, policy enforcement, and interoperability across different deployment models. Addressing these gaps requires collaboration among MNOs, infrastructure vendors, HCPs, standards organizations, and regulatory bodies.

The recommendations outlined below provide strategic guidance to enhance security and operational resilience in 5G cloud environments through a ZT framework.

5.1 Recommendations

5.1.1 For Standards Development and Industry Collaboration

> Align ZTA standards across 3GPP, ETSI, O-RAN, and other bodies to ensure interoperability and a consistent security model.

> Investigate opportunities to standardize data identification and definitions critical for CSM, particularly within the Service-Based Architecture (SBA) and RAN.

> Infrastructure vendors identify opportunities to collaborate with SEIM/SOAR vendors and share threat intelligence with frameworks like MITRE FiGHT and GSMA MOTIF.

> Advance development of 5G-specific policy enforcement architecture for ZTA, including PDP and PEP definitions, use cases, architectures, and situational context for orchestration layers.

5.1.2 For Mobile Network Operators (MNOs) and Infrastructure Vendors

> Adopt a hybrid ZT model supporting both cloud-native and legacy 5G infrastructure with strong IAM, MFA, and continuous authentication.

> Deploy micro-segmentation, service meshes, and eBPF with fine-grained security policy enforcement and enhanced visibility.

  - ○ A thorough risk assessment should be completed before using eBPF because it has the potential of expanding the threat surface.

> Integrate SIEM and SOAR with AI/ML capabilities to detect anomalies, prioritize threats, and automate incident response.

> Plan and execute PQC migration strategies based on NIST, IETF, and ATIS guidance to address future quantum risks [63],[64],[65].

> Support threat and attack data contributions to industry frameworks (MITRE FiGHT, GSMA MOTIF) to strengthen cellular security defenses.

5.1.3 For Cloud Service Providers (CSPs) and Hyperscale Cloud Providers (HCPs)

> Deliver native ZT security-as-a-service solutions for 5G MNOs, including secure container/Kubernetes management and compliance automation.

> Increase transparency by confidentially sharing CVEs and vulnerabilities with 5G MNOs leveraging cloud infrastructure.

> Support Endpoint Detection and Response (EDR) services for private cloud environments, including the option for kernel-based EDR agent deployment in PaaS environments.

> Enhance visibility and reporting on cloud security posture to the MNOs, including controls, logging, configurations, and threat intelligence integration.

5.1.4 For Security Operations and Continuous Monitoring

> Strengthen AI-driven CSM with EDR, anomaly detection, and real-time threat intelligence integration across network layers.

> Ensure ZTA policy enforcement across infrastructure, applications, and data layers through automated DevSecOps pipelines.

> Support AI in SIEM for risk-based alerting, natural language processing for queries, and decision support during incident response.

By implementing these recommendations, 5G MNOs, infrastructure vendors, and HCPs can build a resilient and secure 5G ecosystem, minimizing risks while enabling innovation in next-generation telecommunications infrastructure.

5.2 Next Steps

Subsequent reports in this series of zero trust whitepapers will focus on both practical and forward looking topics related to zero trust policy management in a 5G environment. These include:

- ZTA policy management including functional requirements as well as monitoring and compliance.
- AI/ML Applied to ZT Policy Management
- 5G/6G System Deployment Considerations
- Options for Policy Component Placement and Operation

ACRONYMS AND ABBREVIATIONS

3GPP	3rd Generation Partnership Program
5GC	5G Core
ACME	Automatic Certificate Management Environment
AF	Application Function
AI/ML	Artificial Intelligence and Machine Learning
AMF	Access and Mobility Management Function
API	Application Programming Interface
APT	Advanced Persistent Threat
AUSF	Authentication Server Function
BSS	Business Support Systems
CA	Certificate Authority
CaaS	Container as a Service
CAS	Certificate Authority Service
CI/CD	Continuous Integration and Continuous Delivery/Deployment
CIS	Center for Internet Security
CISA	Cybersecurity and Infrastructure Security Agency
CMP	Certificate Management Protocol
CNF	Container Network Functions
CNI	Container Network Interface
CPU	Central processing unit
CSM	Continuous Security Monitoring
CSR	Certificate Signing Request
CSRIC	Communications, Security, Reliability, and Interoperability Council
CU	Centralized Unit
CVE	Common Vulnerabilities and Exposures
DAST	Dynamic Automated Security Testing
DevSecOps	Development, Security, and Operations
DTLS	Datagram Transport Layer Security
DU	Distributed Unit
eBPF	Extended Berkeley Packet Filters
ECC	Elliptic Curve Cryptography
EDR	Endpoint Detection and Response
ESF	Enduring Security Framework (ESF)
ETSI	European Telecommunications Standards Institute
FCAPS	Fault, Configuration, Accounting, Performance, and Security
FCC	Federal Communications Commission
FIDO	Fast Identity Online

GSMA	Global System for Mobile Communications Association
HCP	Hyperscale Cloud Provider (e.g., AWS, Azure, GCP)
IaC	Infrastructure-as-Code
IaaS	Infrastructure as a Service
IAM	Identity and Access Management
ICAM	Identity, Credentials, and Access Management
IETF	Internet Engineering Task Force
IOC	Indicators of Compromise
KMS	Key Management System
LUKS	Linux Unified Key Setup
M2M	Machine-to-Machine
MANO	Management and Orchestration
MDM	Mobile Device Management
MEC	Multi-Access Edge Computing
MFA	Multi-Factor Authentication
MNO	Mobile Network Operator
MOTIF	Mobile Threat Intelligence Framework
mTLS	Mutual Transport Layer Security
NEF	Network Exposure Function
NF	Network Function
NFV	Network Function Virtualization
NIST	National Institute of Standards and Technology
NRF	Network Repository Function
NWDAF	Network Data Analytics Function
OIDC	OpenID Connect
O-RAN	Open Radio Access Networks
OSS	Operational Support Systems
OTP	One-Time Password
OWASP	Open Web Application Security Project
PaC	Policy-As-Code
PaaS	Platform as a Service
PCF	Policy Control Function
PDP	Policy Decision Point
PEP	Policy Enforcement Point
PIV	Personal Identity Verification
PKI	Public Key Infrastructure
PQC	Post Quantum Cryptography
QoS	Quality of Service
RAN	Radio Access Network
RATS	Remote ATtestation procedureS
RBAC	Role-Based Access Control
RSA	Rivest–Shamir–Adleman

SaaS	Software as a Service
SAML	Security Assertion Markup Language
SAST	Static Automated Security Testing
SBA	Service Based Architecture
SCA	Software Composition Analysis
SCAP	Security Content Automation Protocol
SDO	Standards Development Organization
SEAF	Security Anchor Function
SIEM	Security Information and Event Management
SMF	Session Management Function
SMO	Service Management and Orchestration
SOAR	Security Orchestration, Automation, and Response
SOC	Security Operations Center
SSDF	Secure Software Development Framework
SSH	Secure Shell
SSO	Single Sign-On
STIX	Structured Threat Information eXpression
TAXII	Trusted Automated eXchange of Indicator Information
TDR	Threat Detection and Response
TI	Threat Intelligence
TLS	Transport Layer Security
TOTP	Time-based OTP
UDM	Unified Data Management
UPF	User Plane Function
VM	Virtual Machine
VNF	Virtualized Network Function
WG	Work Group
WIMSE	Workload Identity in Multi System Environments
ZT	Zero Trust
ZTA	Zero Trust Architecture
ZTMM	Zero Trust Maturity Model

REFERENCES

[1] C. Jaikaran, "Salt Typhoon Hacks of Telecommunications Companies and Federal Response Implications," Congressional Research Service, 2025. [Online]. Available: https://www.congress.gov/crs-product/IF12798

[2] Alliance for Telecommunications Industry Solutions (ATIS), "Enhanced Zero Trust and 5G," ATIS White Paper, 2023. [Online]. Available: https://atis.org/resources/enhanced-zero-trust-and-5g/

[3] H. Booth, M. Souppaya, A. Vassilev, M. Ogata, M. Stanley, and K. Scarfone, *Secure Software Development Practices for Generative AI and Dual-Use Foundation Models: An SSDF Community Profile*, National Institute of Standards and Technology (NIST) Special Publication 800-218A, Jul. 2024. [Online]. Available: https://doi.org/10.6028/NIST.SP.800-218A.

[4] S. Rose, O. Borchert, S. Mitchell, and S. Connelly, "Zero Trust Architecture" National Institute of Standards and Technology, Gaithersburg, MD, USA, NIST Special Publication 800-207, Aug. 2020. [Online]. Available: https://csrc.nist.gov/pubs/sp/800/207/final

[5] M. Pope, "Study on Applicability of the Zero trust Security Principles in Mobile Networks," 3rd Generation Partnership Project (3GPP) Technical Report 33.894, v18, 2023. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=4086

[6] M. Pope, "Study on Enablers for Zero Trust Security," 3rd Generation Partnership Project (3GPP) Technical Report 33.794 v.19, 2025. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=4235

[7] 3GPP, *TS 33.502: Securityrelated events handling*, Technical Specification, Draft for Release 20, 2025. [Online] Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=4421

[8] "Network Functions Virtualisation (NFV) Security; Security Management Specification" European Telecommunications Standards Institute (ETSI), 2025. [Online]. Available: https://portal.etsi.org/webapp/workProgram/Report_WorkItem.asp?wki_id=58648

[9] "Multi-access edge Computing (MEC); Study on MEC Security," European Telecommunications Standards Institute (ETSI), GR MEC 041 V3.1.1, 2024. [Online]. Available: https://www.etsi.org/deliver/etsi_gr/MEC/001_099/041/03.01.01_60/gr_MEC041v030101p.pdf

[10] H. Birkholz, D. Thaler, M. Richardson, N. Smith, and W. Pan, "Remote ATestation procedureS (RATS) Architecture," RFC 9334, Jan. 2023. [Online]. Available: https://datatracker.ietf.org/doc/rfc9334/

[11] O-RAN Alliance, "O-RAN: Open Radio Access Network," [Online]. Available: https://www.o-ran.org/.

[12] O-RAN Alliance Security Working Group (WG 11), "Study on Security for O-Cloud," TR.0-R004-v07.00, Oct. 2024. [Online]. Available:https://O-RAN Study on Security for O-Cloud 7.0

[13] O-RAN Alliance Security Working Group (WG 11), "Zero Trust Architecture for Secure O-RAN," May 2024. [Online]. Available: https://mediastorage.o-ran.org/white-papers/O-RAN.WG11.ZTA%20for%20Secure%20O-RAN%20White%20Paper-2024-05.pdf

[14] Cybersecurity and Infrastructure Security Agency (CISA), "Zero Trust Maturity Model," U.S. Department of Homeland Security, April 2023. [Online]. Available: https://www.cisa.gov/zero-trust-maturity-model

[15] O-RAN Alliance Security Working Group (WG 11), "Study on Zero Trust Architecture for O-RAN," 2025. [Online]. Available: https://specifications.o-ran.org/download?id=847

[16] U.S. Department of State, "Announcing the Release of the Administration's National Cybersecurity Strategy," Mar. 2023. [Online]. Available: https://2021-2025.state.gov/announcing-the-release-of-the-administrations-national-cybersecurity-strategy/

[17] Communications Security, Reliability, and Interoperability Council VIII (CSRIC VIII), "Recommendations on the Role of the FCC in Promoting the Availability of Standards for More Secure, Reliable 5G Environment Through the Use of Virtualization Technology," CSRIC VIII Report, June 2023. [Online]. Available: https://www.fcc.gov/sites/default/files/CSRIC8-Report-RecommendationsStandards5GEnvironmentVirtualizationTechnology0623.docx

[18] O-RAN Alliance Cloudification and Orchestration Working Group (WG 6), “Cloud Architecture and Deployment Scenarios for O-RAN Virtualized RAN,” Feb 2025. [Online]. Available: https://specifications.o-ran.org/download?id=817

[19] O-RAN Alliance Use Cases and Overall Architecture Working Group (WG 1), “O-RAN Architecture Description,” Feb. 2025. [Online]. Available: https://specifications.o-ran.org/download?id=789

[20] PCI Security Standards Council, “PCI SSC Cloud Computing Guidelines Version 3,” April 2018. [Online]. Available: https://www.pcisecuritystandards.org/pdfs/PCI_SSC_Cloud_Guidelines_v3.pdf

[21] F. Liu, J. Tong, J. Mao, R. Bohn, J. Messina, L. Badger, and D. Leaf, “NIST Cloud Computing Reference Architecture,” National Institute of Standards and Technology, Gaithersburg, MD, USA, NIST Special Publication 500-292, Sept. 2011. [Online]. Available: https://www.nist.gov/publications/nist-cloud-computing-reference-architecture

[22] Amazon Web Services, “Shared Responsibility Model,” AWS Cloud Security. [Online]. Available: https://aws.amazon.com/compliance/shared-responsibility-model/

[23] Microsoft, “Shared Responsibility in the Cloud,” Azure Security Fundamentals. [Online]. Available: https://learn.microsoft.com/en-us/azure/security/fundamentals/shared-responsibility

[24] Cybersecurity and Infrastructure Security Agency (CISA), “Secure Cloud Business Applications: Hybrid Identity Solutions Guidance,” Secure Cloud Business Applications (SCuBA) Project White Paper, Mar. 2024. [Online]. Available: https://www.cisa.gov/sites/default/files/2024-05/CISA%20SCuBA%20Hybrid%20Identity%20Solutions%20Guidance_0.pdf

[25] Cybersecurity and Infrastructure Security Agency (CISA), “Implementing Phishing-Resistant Multi-Factor Authentication: A CISA Fact Sheet,” Oct. 2022. [Online]. Available: https://www.cisa.gov/sites/default/files/publications/fact-sheet-implementing-phishing-resistant-mfa-508c.pdf

[26] Microsoft, “Multi-factor Authentication,” *Threat Matrix for Kubernetes*. [Online]. Available: https://microsoft.github.io/Threat-Matrix-for-Kubernetes/mitigations/MS-M9001%20Multi-factor%20Authentication/

[27] Open Web Application Security Project (OWASP) Foundation, Kubernetes Security Cheat Sheet,” *OWASP Cheat Sheet Series*, [Online]. Available: https://cheatsheetseries.owasp.org/cheatsheets/Kubernetes_Security_Cheat_Sheet.html

[28] Microsoft, “Microsoft Entra Workload ID,” *Microsoft Security*, [Online]. Available: https://www.microsoft.com/en-us/security/business/identity-access/microsoft-entra-workload-id

[29] Amazon Web Services, Inc., “Multi-Factor Authentication for IAM,” *AWS Identity & Access Management Features*, [Online]. Available: https://aws.amazon.com/iam/features/mfa/#:~:text=AWS%20multi%2Dfactor%20authentication%20(MFA,have%20created%20in%20your%20account

[30] Google Cloud, “Mandatory MFA is Coming to Google Cloud,” *Google Cloud Blog*, Nov. 2024. [Online]. Available: https://cloud.google.com/blog/products/identity-security/mandatory-mfa-is-coming-to-google-cloud-heres-what-you-need-to-know

[31] H. Brockhaus, D. von Oheimb, and S. Fries, “Lightweight Certificate Management Protocol (CMP) Profile”, IETF RFC 9483, Nov. 2023. [Online]. Available: https://datatracker.ietf.org/doc/html/rfc9483

[32] 3rd Generation Partnership Project (3GPP), *Security Architecture and Procedures for 5G System*, Technical Specification 33.501, version 19.2.0, Release 19, Mar. 2025. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3169.

[33] 3GPP, *Network Domain Security (NDS); Authentication Framework (AF)*, Technical Specification 33.310, version 18.5.0, Oct. 2024. [Online]. Available: https://www.3gpp.org/DynaReport/33310.htm.

[34] R. Barnes, J. Hoffman-Andrews, D. McCarney, and J. Kasten, “Automatic Certificate Management Environment (ACME),” *Internet Engineering Task Force (IETF),* RFC 8555, Mar. 2019. [Online]. Available: https://datatracker.ietf.org/doc/html/rfc8555/. [Accessed: Jun. 25, 2025].

[35] 3GPP, *Network Domain Security (NDS); IP Network Layer Security*, Technical Specification 33.210, version 18.0.0, Apr. 2024. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=2279.

[36] International Organization for Standardization (ISO) and International Electrotechnical Commission (IEC), *ISO/IEC 27001:2022 – Information security, cybersecurity and privacy protection – Information security management systems – Requirements*, ISO/IEC, Oct. 2022. [Online]. Available: https://www.iso.org/standard/27001

[37] OWASP Foundation, “OWASP Foundation,” OWASP, 2025. [Online]. Available: https://owasp.org/. [Accessed: Jun. 25, 2025].

[38] Cloud Security Alliance (CSA), “Home | CSA,” [Online]. Available: https://cloudsecurityalliance.org/.

[39] L. Bass and H. T. Morris, “A framework for DevSecOps evolution and achieving continuous integration/continuous delivery (CI/CD) capabilities,” *SEI Blog*, Software Engineering Institute, 21-Apr-2021. [Online]. Available: https://insights.sei.cmu.edu/blog/a-framework-for-devsecops-evolution-and-achieving-continuous-integrationcontinuous-delivery-cicd-capabilities/

[40] The MITRE Corporation, *DevSecOps Best Practices Guide*, version 1.1, Jun. 2023. [Online]. Available: https://saf.mitre.org/DevSecOps_Best_Practices_Guide.pdf

[41] European Union Agency for Cybersecurity (ENISA), *5G Cybersecurity: Cybersecurity Certification in the EU*, Mar. 2022. [Online]. Available: https://www.enisa.europa.eu/publications/5g-cybersecurity-certification-in-the-eu

[42] NIST, *Security Strategies for Microservices-based Application Systems: Addendum to SP 800-204, Volume D*, U.S. Department of Commerce, National Institute of Standards and Technology, Gaithersburg, MD, NIST SP 800-204D, Dec. 2022. [Online]. Available: https://doi.org/10.6028/NIST.SP.800-204D

[43] OWASP Foundation, *CI/CD Security Cheat Sheet*, Open Web Application Security Project, 2023. [Online]. Available: https://cheatsheetseries.owasp.org/cheatsheets/CI-CD_Security_Cheat_Sheet.html

[44] The MITRE Corporation, *“CVE: Common Vulnerabilities and Exposures,”* CVE Program. Page last updated Jun. 4, 2025. [Online]. Available: https://www.cve.org/

[45] Center for Internet Security (CIS), “CIS Benchmarks,” [Online]. Available: https://www.cisecurity.org/cis-benchmarks

[46] National Institute of Standards and Technology (NIST), “Security Content Automation Protocol (SCAP),” NIST Computer Security Resource Center (CSRC) Projects, [Online]. Available: https://csrc.nist.gov/projects/security-content-automation-protocol

[47] Amazon Web Services (AWS), *“Amazon Inspector”*, AWS. [Online]. Available: https://aws.amazon.com/inspector/

[48] Microsoft, *“Microsoft Azure Security Center,”* Microsoft Azure Marketplace. [Online]. Available: https://azuremarketplace.microsoft.com/en-us/marketplace/apps/microsoft.azuresecuritycenter?tab=overview

[49] Google Cloud, *“Security Command Center,”* Google Cloud Security Products. [Online]. Available: https://cloud.google.com/security/products/security-command-center

[50] Google Cloud, *“GKE networking overview,”* Google Kubernetes Engine documentation, last updated Jun. 12, 2025. [Online]. Available: https://cloud.google.com/kubernetes-engine/docs/concepts/network-overview

[51] Amazon Web Services (AWS), *“Alternate CNI plugins for Amazon EKS clusters,”* Amazon EKS User Guide. [Online]. Available: https://docs.aws.amazon.com/eks/latest/userguide/alternate-cni-plugins.html

[52] Microsoft, *“Configure Azure CNI Powered by Cilium in Azure Kubernetes Service (AKS),”* Microsoft Learn, last updated Jun. 13, 2025. [Online]. Available: https://learn.microsoft.com/en-us/azure/aks/azure-cni-powered-by-cilium

[53] Enduring Security Framework (ESF), “Security Guidance for 5G Cloud Infrastructures, Part 1: Prevent and Detect Lateral Movement,” NSA and CISA White Paper, 2021. [Online]. Available: https://media.defense.gov/2021/Oct/28/2002881720/-1/1/0/SECURITY_GUIDANCE_FOR_5G_CLOUD_INFRASTRUCTURES_PART_I_20211028.PDF

[54] Enduring Security Framework (ESF), “Security Guidance for 5G Cloud Infrastructures, Part 2: Securely Isolate Network Resources,” NSA and CISA White Paper, 2021. [Online]. Available: https://www.cisa.gov/sites/default/files/2024-08/SECURITY_GUIDANCE_FOR_5G_CLOUD_INFRASTRUCTURES_PART_II_20211118_508.pdf

[55] 5G Americas, *Advances in Trust and Security in Cellular Wireless Networks in the Age of AI*, Bellevue, WA, Jan. 2025. [Online]. Available: https://www.5gamericas.org/wp-content/uploads/2025/01/Advances-in-Trust-and-Security-AI.pdf

[56] Cybersecurity and Infrastructure Security Agency (CISA), “2023 Top Routinely Exploited Vulnerabilities,” Cybersecurity Advisory AA24-317A, 2024. [Online]. Available: https://www.cisa.gov/news-events/cybersecurity-advisories/aa24-317a

[57] OASIS CTI Technical Committee, “STIX™ Version 2.1 – Core Concepts,” *OASIS Open*, Jul. 2021. [Online]. Available: https://docs.oasis-open.org/cti/stix/v2.1/stix-v2.1.html

[58] OASIS CTI Technical Committee, “Introduction to TAXII,” *OASIS Open*, 2024. [Online]. Available: https://oasis-open.github.io/cti-documentation/taxii/intro.html

[59] MITRE Corporation, "ATT&CK Data & Tools," *MITRE ATT&CK*, [Online]. Available: https://attack.mitre.org/resources/working-with-attack/#accessing-attck-content

[60] MITRE Corporation, "FiGHT™ (5G Hierarchy of Threats)," *MITRE FiGHT™*, 2025. [Online]. Available: https://fight.mitre.org/

[61] GSMA, *FS.57 Mobile Threat Intelligence Framework (MoTIF) Principles*, v1.0, Mar. 2024. [Online]. Available: https://www.gsma.com/solutions-and-impact/technologies/security/wp-content/uploads/2024/04/FS.57-MoTIF-Principles-v1.0.pdf.

[62] 3GPP, *System Architecture for the 5G System (5GS)*, Technical Specification 23.501, version 18.8.0, Jan. 2025. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3144.

[63] U.S. Department of Defense, *The Commercial National Security Algorithm Suite 2.0 – Frequently Asked Questions*, Sep. 2022. [Online]. Available: https://media.defense.gov/2022/Sep/07/2003071836/-1/-1/0/CSI_CNSA_2.0_FAQ_.PDF

[64] National Institute of Standards and Technology (NIST), *Post-Quantum Cryptography Standardization*. [Online]. Available: https://csrc.nist.gov/projects/post-quantum-cryptography/post-quantum-cryptography-standardization

[65] Alliance for Telecommunications Industry Solutions (ATIS), *Preparing 5G for the Quantum Era: An Analysis of 3GPP Architecture and the Transition to Quantum-Resistant Cryptography*. [Online]. Available: https://atis.org/resources/preparing-5g-for-the-quantum-era-an-analysis-of-3gpp-architecture-and-the-transition-to-quantum-resistant-cryptography/

ACKNOWLEDGMENTS

Leadership:
Scott Poretsky, Ericsson
Todd Gibson, T-Mobile
Tom Anderson, ATIS

Participants:
Mike Loushine, AT&T
Dusty Hoffpauir, Charter Communications
Ken Kaminski, Ericsson
Dr. John Srassner, Futurewei Technologies
Dr. Sheeba Backia Mary Baskaran, Motorola Mobility.
Martin Goldberg, National Security Agency
Martin McGrath, Nokia

COPYRIGHT AND DISCLAIMER

ATIS-I-0000104

Published July 2025

Alliance for Telecommunications Industry Solutions
1200 G Street, NW, Suite 500
Washington, DC 20005

The information provided in this document is directed solely to professionals who have the appropriate degree of experience to understand and interpret its contents in accordance with generally accepted engineering or other professional standards and applicable regulations. No recommendation as to products or vendors is made or should be implied.

NO REPRESENTATION OR WARRANTY IS MADE THAT THE INFORMATION IS TECHNICALLY ACCURATE OR SUFFICIENT OR CONFORMS TO ANY STATUTE, GOVERNMENTAL RULE OR REGULATION, AND FURTHER, NO REPRESENTATION OR WARRANTY IS MADE OFMERCHANTABILITY OR FITNESS FOR ANY PARTICULAR PURPOSE OR AGAINST INFRINGEMENT OF INTELLECTUAL PROPERTY RIGHTS. ATIS SHALL NOT BE LIABLE, BEYOND THE AMOUNT OF ANY SUM RECEIVED IN PAYMENT BY ATIS FOR THIS DOCUMENT, AND IN NO EVENT SHALL ATIS BE LIABLE FOR LOST PROFITS OR OTHER INCIDENTAL OR CONSEQUENTIAL DAMAGES. ATIS EXPRESSLY ADVISES THAT ANY AND ALL USE OF OR RELIANCE UPON THE INFORMATION PROVIDED IN THIS DOCUMENT IS AT THE RISK OF THE USER.

NOTE - The user’s attention is called to the possibility that compliance with this standard may require use of an invention covered by patent rights. By publication of this standard, no position is taken with respect to whether use of an invention covered by patent rights will be required, and if any such use is required no position is taken regarding the validity of this claim or any patent rights in connection therewith. Please refer to [http://www.atis.org/legal/patentinfo.asp] to determine if any statement has been filed by a patent holder indicating a willingness to grant a license either without compensation or on reasonable and non-discriminatory terms and conditions to applicants desiring to obtain a license.

www.atis.org

For information, contact ATIS at (202) 628-6380.