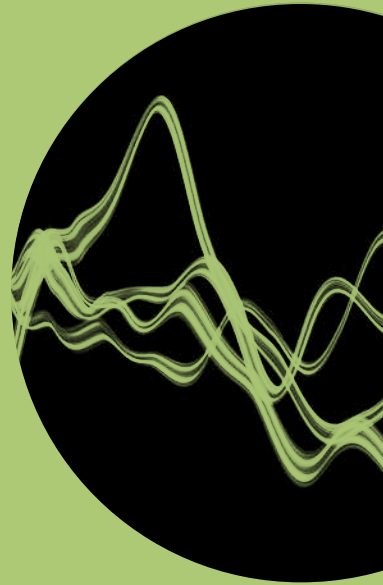


Review

ERICSSON
TECHNOLOGY



REALIZING
ZERO TRUST
IN 5G NETWORKS



ERICSSON

5G zero trust

– A ZERO-TRUST ARCHITECTURE FOR TELECOM

The heterogeneous nature of modern telecommunications infrastructure is making it increasingly difficult to protect network resources with conventional perimeter-oriented approaches to network security. By starting from the assumption that the attacker is already inside the network, the zero trust model enhances security by both blocking unauthorized access to network resources and preventing internal lateral movement by an attacker.

JONATHAN OLSSON,
ANDREY SHOROV,
LOAY ABDELRAZEK,
JORDEN WHITEFIELD

The primary aim of any approach to network security is to protect the communication infrastructure so that it can provide services with the expected level of quality, free of disruption. By significantly mitigating risks inside the network perimeter, the zero trust model makes it easier for communication service providers (CSPs) to live up to their security commitments.

■ The perimeter security model operates on the basis of inherent trust, assuming that everything on the inside of a network is trustworthy. As long as the attacker is outside the network and the outer perimeter defenses are strong enough to completely prevent breaches, this approach can work well. But if

a breach does occur and an attacker gets inside the network, the perimeter security model allows the attacker to move laterally between systems within the network.

The zero trust (ZT) security model resolves this issue by never making any assumptions about trustworthiness. It first emerged more than a decade ago in the enterprise space, which means the telecommunications sector benefits from the enterprise sector's findings and best practices.

A zero trust architecture (ZTA) works by facilitating secure network access to resources (data, devices and services) that is limited only to subjects (users, devices and services) that are authorized and approved. It is built on an identity-centric approach based on the execution of policy-based authorization

decisions in runtime combined with traditional defense-in-depth security principles. When implemented correctly, a ZTA mitigates both the risk of an external attacker getting a foothold in the network as well as the risk of lateral movement, in the case of a security breach.

Ericsson's approach to zero trust architecture applies the ZT principles [1, 2] to telecommunications networks. We have chosen to use the terminology and tenets defined by the US National Institute of Standards and Technology (NIST) SP 800-207 [3] (see highlight box). Several other government bodies and organizations are, however, in the

process of publishing ZTA guidance or requirements. The National Cyber Security Centre in the United Kingdom currently has its own ZTA design principles [4]. The NSA and US Cybersecurity and Infrastructure Security Agency's Trusted Internet Connections initiative [5, 6, 7] also aligns with ZT principles.

Built-in support for zero trust architecture in 5G

The 3GPP 5G standards define relevant network security features supporting a zero trust approach in the three domains: network access security, network domain security and service-based architecture (SBA) domain security.

Seven tenets for zero trust architecture

The US National Institute of Standards and Technology has defined seven tenets for zero trust architecture [3]:

T1. All data sources and computing services are considered resources. Devices in a network are heterogeneous and they all interact with the network and software services.

T2. All communication is secured regardless of network location. Trust of a device based on where it is located in a network is not enough. All communication should be secure – that is, confidentiality and integrity must be maintained.

T3. Access to individual [operator] resources is granted on a per-session basis. Trust of devices and services is evaluated prior to granting access. Access is ephemeral and only the minimal set of privileges required are granted for the session.

T4. Access to resources is determined by dynamic policy – including the observable state of client identity, application/service, and the requesting asset – and may include other behavioral and environmental attributes. Clients accessing resources are granted access and permissions based on the client's ascertained state and access rules defined in policies.

T5. The [operator] monitors and measures the integrity and security posture of all owned and associated assets. Trust nothing, verify everything. When a request to access a resource appears, the asset is evaluated. The evaluation of assets is continuous, so as to have an accurate assessment of the threat landscape and risks.

T6. All resource authentication and authorization are dynamic and strictly enforced before access is allowed. Authentication and authorization are always required before accessing any resource for a limited time period and this is continual – that is, reauthentication and reauthorization occurs throughout all transactions where required.

T7. The [operator] collects as much information as possible about the current state of assets, network infrastructure and communications and uses it to improve its security posture. Collected data provide context and insights about where security improvements are needed, such as evaluating access requests, optimizing policy creation and enforcement.

THE ABILITY TO SECURELY PROVISION, STORE, ACCESS, USE AND REVOKE CREDENTIALS IMPACTS TRUSTWORTHINESS

The network access security features provide users with secure access to services through the device (mobile phone or connected IoT device) and protect against attacks on the air interface between the device and the radio node.

Network domain security includes features that enable nodes to securely exchange signaling data and user data, for example, between radio and core network functions (NFs) [8].

The 5G SBA is built on web technology and web protocols to enable flexible and scalable deployments using virtualization and container technologies and cloud-based processing platforms. SBA domain security specifies the mechanism for secure communication between NFs within the serving network domain and with other network domains.

Key 5G security features that enable zero trust architecture

In our assessment, there are four key security features in 5G that are of most significance in terms of enabling zero trust architectures: secure digital identities, secure transport, policy frameworks and security monitoring.

Secure digital identities

Identities are the new perimeter to defend in ZT security, as they are the primary factor that determines whether access to resources is granted. Secure digital identities consist of two parts. The first part is the identifier (username, fully qualified domain name, serial number) that uniquely identifies a subject or resource. The second part is the credential (password, private key, token) that is secret data used to verify the authenticity of the subject or resource. The use of secure digital

identities must be complemented with processes and technologies that enable the secure management of identities and credentials.

In 5G, each and every subject (subscriber or gNodeB, for example) and resource (such as an SBA NF) is uniquely identifiable. Secure digital identities play a fundamental role in building trust and securing communication between entities across security domains. Examples include the digital identity in SIM cards used to authenticate subscribers and network access control, digital identities based on X.509 certificates used for mutual authentication of network devices and NFs, and management user identities for management access control. Secure digital identities enable the creation of an inventory of network assets and are critical to enabling the authentication of subjects and resources to satisfy NIST tenets T2, T3, T4 and T6.

Confidence in the trustworthiness of an identity is determined by the ability to authenticate the asserted identity and the ability to ascertain the integrity of the device being authenticated. The ability to securely provision, store, access, use, renew and revoke credentials impacts trustworthiness.

Identity life cycle management is more challenging for virtual network functions (VNFs) than it is for network appliances. Firstly, the dynamic nature of virtualized deployments – where NFs are instantiated and removed depending on demand – requires secure provisioning, removal and revocation of digital identities in multi-tenant environments. Secondly, consistent exposure and availability of secure hardware across cloud platforms is needed for secure storage and limited access to key material. Secure hardware is used to protect against theft and misuse of secrets, particularly in multi-tenant environments. Further development and maturity in cloud deployments will be required to protect digital identities and attest the system integrity in 5G networks.

Secure transport

The T2 requirement that all communications must be secured is aligned with 3GPP 5G standards that are developed under the presumption of open

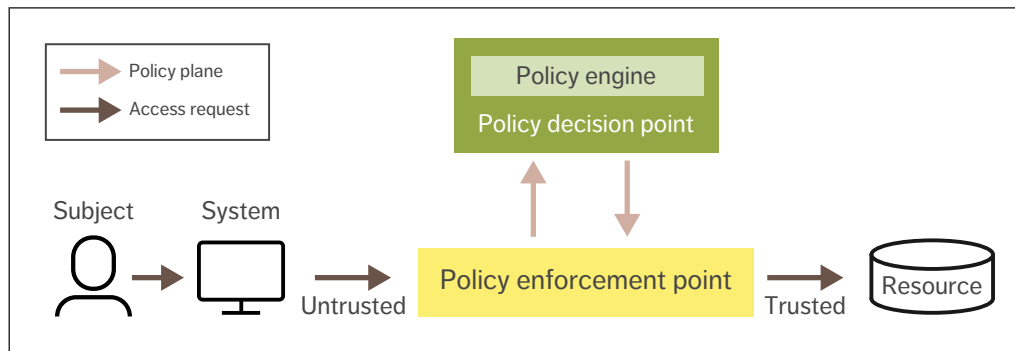


Figure 1 The logical components of the policy framework

networks in which all links could be intercepted. Industry-standard mechanisms are used to secure the communication of user and signaling data across 3GPP interfaces.

Data between the user equipment and the radio base station is secured with cryptographic algorithms, providing confidentiality and integrity protection. Additional improvements are introduced in 5G with the Subscription Concealed Identifier (SUCI) to further enhance protection of subscriber privacy against conventional attacks such as passive eavesdropping or active probing of permanent and temporary identifiers.

Communication in transport networks, and between NFs and interconnect networks, is secured with industry-standard security protocols such as (D)TLS 1.2 and 1.3, IPsec, and MACsec, all of which support mutual authentication.

Policy frameworks

The relationships and interactions between the many logical and physical entities in telecommunication networks must be managed to ensure that resources are only accessed by authorized subjects. Policies capture the access rules and requirements to determine the eligibility of a request. These policies are managed, distributed and enforced by a policy framework [3, 9]. This enables the enforcement of micro-perimeters with fine-grained access control based on roles, credentials and environmental attributes.

Figure 1 presents the logical components of a policy framework. The essential logical entities are the policy decision point (PDP) and policy enforcement point (PEP). To access the specific resource, a subject requests permission from the PDP and provides the information needed to perform authentication and authorization.

Policies are created to reflect an organization's processes and acceptable level of risk as well as the sensitivity of the targeted asset. A policy specifies the required level of protection for an object, privileges of a subject and environmental conditions that can change the allowed behavior of the subject toward the object.

The policy engine is part of the PDP. It runs a trust evaluation algorithm to calculate a subject's trust score, which is used to determine whether the subject is allowed to access a resource. The trust algorithm may only use information provided by the subject or it may utilize additional metadata (geographic location of the subject, historical resource usage and behavior).

The PEP is a component that is responsible for setting up a micro-perimeter to protect a resource. Where possible, the PEP is integrated into the resource or placed as close as possible to it, and it forms a logical demarcation point between security zones. The PEP provides access control of connections between the subject and resource based on access control decisions from the PDP.

Policy frameworks are employed in 3GPP-based systems to manage access to resources in different security domains. For example, to gain access to the 5G network services (T1), the user equipment (UE) contacts an Access and Mobility Management Function (AMF) that takes a PEP role. A PDP role can be represented by multiple NFs where Unified Data Management (UDM) and the Policy Control Function (PCF) may be highlighted, among others.

THE SECURITY POSTURE OF THE REQUESTING ENTITY MUST BE EVALUATED BY DYNAMIC ACCESS CONTROL POLICIES

The AMF transmits the UE's access request to the UDM to validate the UE's identity and trigger authentication and authorization procedures to establish a secure channel (T2, T6). The PCF feeds the AMF with access and mobility policies that may affect UE authorization to access 5G network resources due to, for example, mobility restrictions (T4) [10, 11, 12, 13].

Another example describes how ZT principles apply in 5G SBA. In reference to T1, the SBA identifies NF service consumers and NF service producers. Communication security between core NFs has improved significantly in comparison with previous generations of mobile networks. SBA security specification requires the performance of Transport Layer Security (TLS) based mutual authentication and OAuth 2.0 token-based authorization for any NF that wants to communicate with another NF (T2, T6). The network repository function (NRF) takes the role of authorization server, which makes the NRF act as the PDP.

The introduction of the service communication proxy (SCP) allows indirect communication between NFs. The SCP can take the role of the PEP and provide access control functionality by requesting authorization decisions from the NRF.

This makes it possible to implement the zero trust model in the 5G Core, where an NF service consumer (subject) requests access to an NF service producer (resource) through the SCP (PEP), and the NRF (PDP) grants or denies access [10, 13, 14]. With regard to T4, to support decision-making about requested access to resources, the NRF can store additional information, defining the actions allowed for an NF service consumer to specific NF producers [13].

Security monitoring

Security monitoring supports the detection of threats and measuring the security posture of network assets and compliance with security policies. Monitoring and evaluation of subjects, resources compliance, trustworthiness and state are important when deciding whether to permit access to resources.

The European Telecommunications Standards Institute (ETSI) defines security and trust guidance for NFs [15]. With guidelines emphasizing that compliance and state measurements must be continually monitored to effectively evaluate the level of trust of an NF, ETSI's guidance adheres with the principles of zero trust design.

In line with T3 and T4, the security posture of the requesting entity must be evaluated by dynamic access control policies before access is granted to the requested resource. Additionally, to satisfy T5, all owned assets in a telecom network should be monitored and their security posture should be evaluated continuously. These assets include, but are not limited to, devices accessing the network, RAN NFs, core NFs and management functions.

There are different ways to implement a trust evaluation algorithm. Identifying which trust algorithm implementation to adopt depends on its characteristics:

1. How different parameters are evaluated (as binary decisions or as weighted parts of a whole score or confidence level)
2. How requests are evaluated in relation to other historical requests by the same subject.

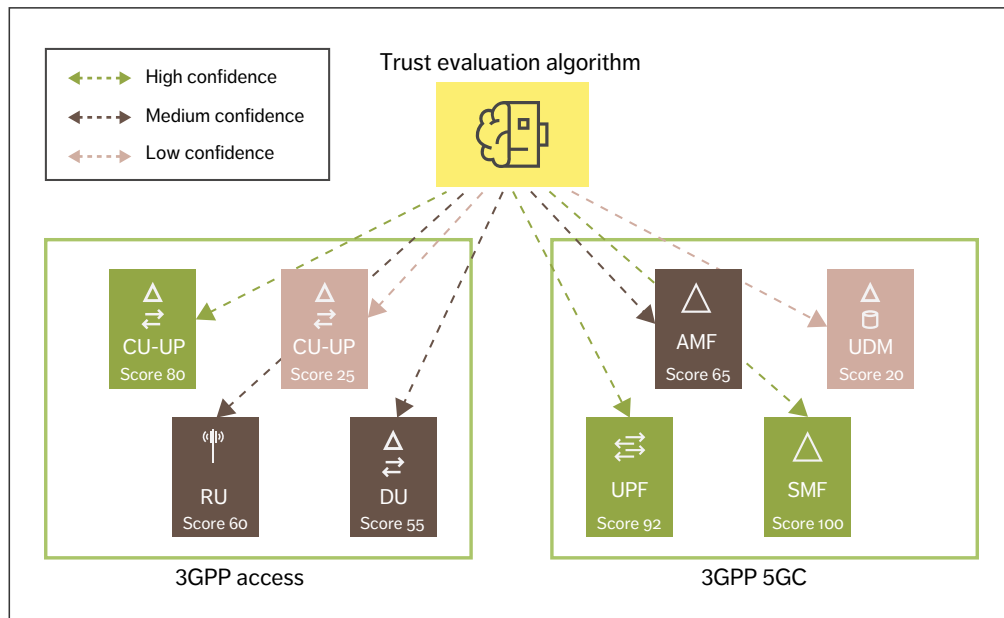


Figure 2 NF trust evaluation

Parameters can be evaluated either based on criteria or score [3]. Score-based evaluation computes a confidence level based on values from every data source, recognizing that there may be various levels of trust between different subjects. Criteria-based evaluation relies on a set of statically configured attributes that must be met before access is granted to a resource or an action is allowed. Moreover, requests can be evaluated either singularly or contextually. Singular evaluation treats each request individually, which risks that an attack can go undetected. Unlike singular evaluations, contextual evaluation takes the subject's history into consideration when evaluating access requests.

The implementation of a trust evaluation algorithm that combines contextual, score-based characteristics would make it possible to offer dynamic and granular access control, since the score provides a confidence level for the requesting account and adapts to changing factors more quickly than static policies.

With respect to T7, there are multiple parameters [15] that can be taken into consideration for evaluating trust that are relevant for telecommunication networks. Examples include geographical location, NF location, software capabilities (such as patch level, software versions), execution history of an instance, configuration compliance and the appropriate use of encryption techniques.

Future telecommunication networks should not only consider how to handle trust in subjects, but also trust in resources – particularly in multi-vendor deployments or in cloud where services are provided by a third party.

Figure 2 illustrates how each NF instance may have different trust scores based parameters such as on their current configured state, image version and compliance level. The trust evaluation algorithm in Figure 2 assigns scores in three different ranges on a scale of 0-100: low confidence (<50), medium confidence (51-79) and high confidence (>=80). Based on the evaluation at a certain point in time,

●● THE TRANSITION TOWARD ZERO TRUST REPRESENTS A MAJOR STEP CHANGE FOR THE TELECOM INDUSTRY ●●

a score is provided to the NF. If the score falls below a certain threshold, further actions should be taken, such as terminating the NF and replacing it.

Next steps

Telecommunications standards have already evolved the telecom security model by adopting ZT principles that better reflect the security reality facing CSPs. While the latest standards provide improved management of security risks for robust and reliable networks and services, a CSP's ability to fully implement a zero trust architecture will also depend on additional technologies, prioritizations and processes.

Consequently the journey towards zero trust will be gradual with methodical decisions on when, where and how to deploy new security technologies and processes. One of the first important decisions a CSP needs to make is whether the transition should include existing infrastructure and, if so, how to include it with minimal operational and security risk. For example, traditional controls should not be decommissioned until careful evaluation and testing of the new security controls has been completed.

The gradual introduction of zero trust principles, process changes and technology solutions should be driven by risk-based decisions about when and where a CSP wants to invest in modernizing its technologies and business processes. Future challenges include the need to manage the risks of both the infrastructure that has migrated to ZTA and the infrastructure that has not.

A successful implementation of ZT builds on the foundation of effective information security and resiliency practices. While a ZTA can help focus security efforts, it is not by itself sufficient to realize a

secure architecture. Rather, a ZTA serves as a cornerstone of a holistic active defense strategy for managing risk, complementing established state-of-the-art information security practices.

Today's concept of zero trust, which focuses on network security, will need to evolve in the years ahead. It will need to expand to tackle the issue of how to address vertical trust from the application, the execution environment and device hardware in cloud environments. This includes measuring the system when instantiating network functions and determining the integrity and origin of software.

Additionally, confidential computing technologies to protect software and data will be critical to protect sensitive assets in shared and distributed environments. Hardware rooted security will be essential to establish a verifiable chain of trust from the hardware to the applications that run on it, as well as protecting data in transit, at rest and in use, to address the risks introduced by hardware and software disaggregation and multivendor deployments.

All of these various technical challenges require further research, development and standardization to fully realize the potential of ZTA for the telecom industry.

Conclusion

The transition toward zero trust represents a major step change for the telecom industry. Ericsson is committed to delivering solutions that enable communication service providers (CSPs) to make that transition as smooth as possible. Fortunately, the new requirements and functionalities introduced in the 5G specifications are already aligned with many of the zero trust tenets. It is already evident, however, that further technology development, standardization and implementation are needed in areas such as policy frameworks, security monitoring and trust evaluation to support the adoption of zero trust architecture in new telecom environments that are distributed, open, multi-vendor and/or virtualized.

While various technologies can support organizations in adhering to the guiding principles of

zero trust as part of their total active defense strategy, it is important to remember that technology alone will never be sufficient to realize the full potential of zero trust. Successful implementation of a network based on zero trust principles requires the concurrent implementation of information security processes, policies and best practices, as well as the presence of knowledgeable security staff. Regardless of where a CSP is in its transition toward a zero trust architecture, the three pillars of people, processes and technology will continue to be the foundation of a robust security architecture.

●● TECHNOLOGY
ALONE WILL NEVER BE
SUFFICIENT TO REALIZE
THE FULL POTENTIAL OF
ZERO TRUST ●●

Terms and abbreviations

AMF – Access and Mobility Management Function | **CSP** – Communication Service Provider | **CU-UP** – Central Unit User Plane | **DU** – Distributed Unit | **ETSI** – European Telecommunications Standards Institute | **NF** – Network Function | **NIST** – National Institute of Standards and Technology | **NRF** – Network Repository Function | **PCF** – Policy Control Function | **PDP** – Policy Decision Point | **PEP** – Policy Enforcement Point | **RU** – Radio Unit | **SBA** – Service-Based Architecture | **SCP** – Service Communication Proxy | **SMF** – Session Management Function | **SUCI** – Subscription Concealed Identifier | **TLS** – Transport Layer Security | **UDM** – Unified Data Management | **UE** – User Equipment | **UPF** – User Plane Function | **VNF** – Virtual Network Function | **ZT** – Zero Trust | **ZTA** – Zero Trust Architecture

Further reading

- » **Ericsson, 3GPP 5G security overview, available at:** <https://www.ericsson.com/en/blog/2019/7/3gpp-5g-security-overview>
- » **Ericsson white paper, Building trustworthiness into future mobile networks, available at:** <https://www.ericsson.com/en/reports-and-papers/white-papers/building-trustworthiness-into-future-mobile-networks>
- » **Ericsson, Future network security, available at:** <https://www.ericsson.com/en/future-technologies/future-network-security>
- » **Ericsson, Security, available at:** <https://www.ericsson.com/en/security>

References

1. O'Reilly Media, Inc., **Zero Trust Networks: Building Secure Systems in Untrusted Networks**, first edition, 2017, Gilman, E; Barth, D
2. Gartner, **Market Guide for Zero Trust Network Access**, June 8, 2020 (retrieved December 1, 2020) Riley, S; MacDonald, N; Orans, L, available from: <https://www.gartner.com/en/documents/3986053/market-guide-for-zero-trust-network-access>
3. National Institute of Standards and Technology, **NIST SP 800-207 Zero Trust Architecture**, August 11, 2020 (retrieved December 1, 2020), Rose, S; Borchert, O; Mitchell, S; Connelly, S; available from: <https://csrc.nist.gov/publications/detail/sp/800-207/final>
4. UK NCSC, **Zero trust architecture design principles**, available at: <https://github.com/ukncsc/zero-trust-architecture/>
5. Cybersecurity and Infrastructure Security Agency, **Trusted Internet Connections 3.0 – TIC Core Guidance Volume 2: Reference Architecture**, July 2020, available at: https://www.cisa.gov/sites/default/files/publications/CISA_TIC%203.0%20Vol.%202%20Reference%20Architecture.pdf
6. Cybersecurity and Infrastructure Security Agency, **Trusted Internet Connections 3.0 – TIC Core Guidance Volume 3: Security Capabilities Catalog**, July 2020, available at: https://www.cisa.gov/sites/default/files/publications/CISA_TIC%203.0%20Vol.%203%20Security%20Capabilities%20Catalog.pdf
7. National Security Agency Central Security Service, **Zero Trust Security Model**, February 26, 2021, available at: https://media.defense.gov/2021/Feb/25/2002588479/-1/-1/0/CSI_EMBRACING_ZT_SECURITY_MODEL_UOO115131-21.PDF
8. Ericsson, **An overview of the 3GPP 5G security standard**, July 17, 2019, Ben Henda, N; Wifvesson, M; Jost, C, available at: <https://www.ericsson.com/en/blog/2019/7/3gpp-5g-security-overview>
9. National Institute of Standards and Technology, **NIST special publication 800-162, Guide to Attribute Based Access Control (ABAC) Definition and Considerations**, (updated) August 2, 2019 (retrieved December 1, 2020), Hu, C. T; Ferraiolo, D.F; Kuhn, D.R; Schnitzer, A; Sandlin, K; Miller, R; Scarfone, K, available from: <https://www.nist.gov/publications/guide-attribute-based-access-control-abac-definition-and-considerations-0>
10. **3GPP TS 23.501: System Architecture for the 5G System**, Release 16, available at: <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3144>
11. **3GPP TS 23.502: Procedures for the 5G System**, Release 16, available at: <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3145>
12. **3GPP TS 23.503: Policy and charging control framework for the 5G System**, Release 16, available at: <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3334>
13. **3GPP TS 33.501: Security architecture and procedures for the 5G system**, Release 16, available at: <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3169>
14. Ericsson, **Security for 5G Service-Based Architecture: What you need to know**, August 21, 2020, Jost, C; Smeets, B, available at: <https://www.ericsson.com/en/blog/2020/8/security-for-5g-service-based-architecture>
15. ETSI **GS NFV-SEC 003, Network Functions Virtualisation (NFV); NFV Security; Security and Trust Guidance**, December 2014, available at: https://www.etsi.org/deliver/etsi_gs/NFV-SEC/001_099/003/01.01.01_60/gs_NFV-SEC003v010101p.pdf

THE AUTHORS



Loay Abdelrazek

◆ joined Ericsson in 2019 as a researcher focusing on security concepts in RAN. His research explores new security concepts and technology for RAN, including topics such as air interface security, cloud RAN security and systems security. He holds an M.S. in cybersecurity from Nile University, Giza, Egypt.

Jonathan Olsson

◆ joined Ericsson in 2004 as a researcher for fixed access networks. Since then, he has held roles

including standardization coordinator, strategic product manager and security leader in the CTO office. In his current role as RAN security leader, Olsson drives RAN security technology strategy activities in areas such as cloud security, intrusion detection and response, Internet of Things security and trust technologies. He has a B.Sc. in computer science from Uppsala University, Sweden, and is a certified information systems security professional.



Andrey Shorov

◆ is a specialist in security technology at Ericsson Network Security who joined the company in 2019. He identifies key security technologies for the 5G network infrastructure and network slicing. Shorov holds a Ph.D. in computer science from the St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences.

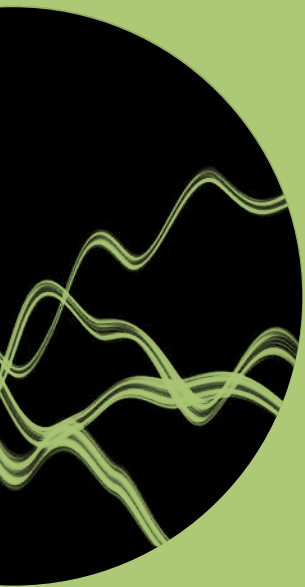
Jorden Whitefield

◆ has served as a security technology specialist at

Ericsson Network Security since 2019. As an ethical hacker, he performs product security testing on emerging 5G mobile network products, with a focus on platform and operating system security. Whitefield holds a Ph.D. in computer science from the University of Surrey in Guildford in the UK. His doctoral thesis was on the subject of formal verification of security protocols.



The authors would like to thank Ari Pietikäinen, János Köver, Patrik Teppo, Ilhan Gurel, Mathias Weibull and Antti Jaakkola for their valuable contributions to this article.



ISSN 0014-0171
284 23- 3358 | Uen

© Ericsson AB 2021
Ericsson
SE-164 83 Stockholm, Sweden
Phone: +46 10 719 0000