A guide to 5G network security 2.0

Conceptualizing security in mobile communication networks – how does 5G fit in?

ERICSSON


Username
New policy set
Duplicate
Delete
State
Assets
Controls
Compliance status
Scheduled checks
Last configured
In production
In testing
Create new asset
Asset vMTAS
1. Properties
2. Select asset type
3. Connectivity
4. Bind with policy sets
Set the connectivity for the asset
Step 3 of 4
Asset type
vMTAS 1.6
Asset interfaces*
Netconf (Netconf)
Asset components
SC1 (SC)
SC2 (SC)
PL1 (PL)
PL2 (PL)
Interface content
Name*
Netconf
IP address *
172.221.53.91
Strict Host Key Checking
Yes
No
Authentication method
Username
SSH private key
Privileged escalation method
Sudo
Su
Shared jumphost*
Yes
No
Jumphost
Select
Port
830
SSH host key
Cancel
Save and create a new asset
Save and exit
Previous

Executive summary

An introduction to telecommunication network security

- Today's mobile telecommunication networks are generally separated into four logical parts: radio access network, core network, transport network and interconnect network. Each network part comprises three so-called planes, each of which is responsible for carrying a different type of traffic, namely: the control plane which carries the signaling traffic; the user plane which carries the payload (actual-) traffic; and the management plane which carries the management traffic. In terms of network security, all three planes can each be exposed to unique types of threats. There are also uniform threats which can affect all three planes simultaneously.
- Telecommunication network security is defined by the following layers that determine the network security experience of end users:
  - **Network operation**: the operational processes which allow networks to function and deliver targeted levels of security are highly dependent on the deployment and operations of the network itself.
  - **Network deployment**: at the deployment phase, networks are configured for a targeted security level, which is key to setting security parameters and further strengthening the security and resilience of the network.
  - **Vendor product development**: network vendors design, develop and implement the agreed standards for functional network elements and systems, which play a crucial part in making the end network product both functional and secure.
  - **Telecommunication standardization**: a process whereby operators, vendors and other stakeholders set standards for how networks around the globe will work together. This also includes how best to protect networks and users against malicious actors.

In qualitative terms alone, 5G is worlds ahead of 4G

- From a user perspective, 5G is inherently different to any of the previous mobile generations. Machine-type communication, enabled by 5G, is widely anticipated to become the strategic difference and unique selling point of 5G in the long run. 5G networks will serve as critical infrastructures to facilitate the digitization, automation, and connectivity to machines, robots, and transport solutions etc. Thus, there is a significant value at stake and, so too, a significantly different tolerance for risk.
- 5G marks the beginning of a new era of network security with, for example, the introduction of IMSI encryption. Bodies, such as those represented through 3GPP, do not standardize how functions are implemented and realized, this is done by the vendors and hence the standard as such is only one aspect of the security posture of a deployed network. The main purpose of the specifications is to secure interoperability between the functions required to provide network connectivity. Consequently, there is little about virtualization and cloud deployments in the specifications. These details will be addressed at the implementation and deployment phases.[1]

Understanding security in the era of 5G

Telecom networks are evolving rapidly across a broad technological environment which includes virtualization, disaggregation, cloud, AI, IoT and Industry 4.0. This is met by an equally broad yet an increasingly challenging cybersecurity environment.

Advances in technology, together with the broader development of networks beyond 5G radio access network (RAN), are expected to have a significant impact on security, such as for software-defined networking (SDN), network function virtualization (NFV) and edge computing. The 5G 3GPP standard is agnostic, in that it is flexible enough to allow for different types of physical and virtual overlap between the radio access network (RAN) and core network, for example, from a remote device to the core network. The separation and deployment of functions between RAN and core raises questions about competitiveness and performance. From an economic, competitive and performance perspective, failing to make use of technological developments in the configuration and deployment of 5G commercial networks will ultimately prove counterproductive to realizing unique 5G use cases, such as critical machine-type communication or applications which belong to latency-sensitive autonomous systems.

The disaggregation of RAN also comes with the challenge, albeit not unsurmountable, of securely integrating the disaggregated O-RAN solution. As of 2021, secure O-RAN solutions will require additional security measures that are not fully addressed by 3GPP SA3 security standards, since O-RAN introduces additional open interfaces and functions (such as lower-layer split and near Real Time RIC) that are not part of the 3GPP standard.[2]

In the era of 5G, it's important that, when we begin to conceptualize security on a system wide level where telecom networks are an important component, we adopt a strong understanding of the following:

- Increased value at stake and decreased risk tolerance
- Cyber-physical dependencies
- Security of standards, products, deployments, and operations
- Proactive cybersecurity measures
- Vulnerability management
- Securing the supply chain

Contents

1. Introduction

Enhancements in mobile telecommunication networks are galvanizing a wave of digital transformation which is disrupting industries of all types and forcing us to rethink our traditional ways of working. This transformation is not just changing how we work with IT, office tools and administrative systems; but it's also creating new business opportunities. Value chains are becoming value networks, where one-to-one relations between suppliers, vendors, operators, and end users are being reinvented as ecosystems of partners and co-creators.

Internet of Things and Industry 4.0
This cross-industry transformation has created a need to evolve the concept of wireless connectivity for the fifth generation of mobile technology (5G[3]), to enable new ways of defining performance monitoring and assurance as well as quality of service and user experience.[4] Compared with previous generations of wireless communications technology, the rationale for 5G development is to expand the broadband capability of mobile networks to provide specific new value propositions not only for consumers but also for various industries and society at large. Hence, unleashing the potential of the new consumer use cases as well us industrial use cases including massive deployments of Internet of Things[5] (IoT).

With IoT and Industry 4.0, a plethora of new device types with less homogeneity than today's PCs and smartphones will be connected with new and broader sets of applications. Not just internet-based apps and content, but rather real time, mission-critical, industrial control systems (Supervisory Control and Data Acquisition, -SCADA) systems. The next digital era will not just be confined to data behind screens and keyboards but will also enter the cyber-physical domain through robots, sensors, and autonomous cyber-physical processes.

Digital transformation will further introduce new dimensions of attack vectors, values, and vulnerabilities through these connected digital systems. IoT brings a new set of issues, such as the security, safety, and robustness of cyber-physical systems. Novel types of attack, as well as new privacy and cybersecurity regulations, may take many industries by surprise.

Cloud and Mobile Edge Computing
Edge computing is a form of cloud computing that pushes the data processing power out to the edge devices rather than centralizing compute and storage in a single data center. This is a distributed computing paradigm that brings computation and data storage closer to the location where it is needed. This reduces latency and network contention between the equipment and the user, which increases responsiveness. Efficiency may also improve because only the results of the data processing need to be transported over networks, which consumes far less network bandwidth than traditional cloud computing architectures.

Edge computing in telecom, often referred to as Mobile Edge Computing or Multi-Access Edge Computing (MEC), provides execution resources for applications with networking close to the end users, typically within or at the boundary of operator networks. This opens up a new business ecosystem and value chain in the telecom industry. Operators can now open up their Radio Access Network (RAN) edge nodes to authorized third parties to deploy their innovative applications flexibly and rapidly at the telco edge to serve their customer better.

While MEC brings new opportunities, technological advancements, new business models, new ecosystem partners into the telecom world, it also brings additional security challenges to protect both the operator and the third-party resources. Security becomes one of the key issues as already seen in a cloud computing world. Edge computing, being close to the end user, is more susceptible to physical attacks compared to traditional data center-centric cloud computing. In general, the same type of attacks is applicable to edge computing as in cloud computing, but the attack surface gets much bigger. This also means that better security monitoring and management capabilities at the edge needs to be developed.

AI, ML and automation
Artificial Intelligence (AI) refers to computer systems able to perform tasks that typically require human intelligence. AI is not a specific method but rather a set of approaches and techniques, including Machine Learning (ML), machine reasoning, and robotics. Specifically, ML refers to components of AI systems that learn from data to make predictions or decisions without being explicitly programmed to perform specific tasks like classification, regression, anomaly detection, or clustering.

AI/ML technologies[29] are being used widely in telecom products and services at all layers of the network. They play a pivotal role in digital transformation through automated decision-making capabilities. The high bandwidth, low latency, and massive connectivity of 5G systems requires a high degree of security automation, which cannot be achieved without AI/ML. AI/ML brings excellent opportunities to improve security solutions by supporting and automating security operations, automating incident analysis, finding, and predicting security breaches, and dynamically adapting to changing conditions. Ericsson is actively increasing efforts to apply AI/ML for telecom security. AI/ML capabilities are presented in a wide variety of telecom security use cases, starting from detecting fraudulent usage of mobile subscriptions or spotting false base stations to cell tower inspection.

However, AI/ML and security is a two-edged sword. While it enables more powerful security solutions, it also presents opportunities for new types of attacks that are unique to AI/ML. These attacks include techniques such as subverting system operation through small, subtle changes in

input data (model evasion attacks); learning, or stealing, the models (model extraction attacks); and inferring facts about the data used to build the model (model inversion attacks). Because AI/ML depends heavily on data both during development and deployment, data security and privacy is a key concern. These attacks, and potential defenses, are an extremely dynamic area of cutting-edge research. Ericsson is actively involved in driving that research and applying it to telecom systems, which will be key to establishing trustworthy AI/ML.

Open Source Software[6]

The utilization of Free Open Source Software (FOSS) has many positives for ICT industry. Depending on the FOSS scope and maturity, these may include R&D cost reduction, agility, productivity increase, source code auditability and fast bug fixing. Large development communities could benefit from passionate individuals with expert knowledge that would be hard to in-source.

Improved security is often quoted as a major benefit, that is however not automatically the case. When the FOSS code base grows, it is increasingly hard to review and test all the code for potential vulnerabilities. There are ample examples of both unintentional and purpose-built vulnerabilities in widely used FOSS. Some of the vulnerabilities have remained undetected for years which has given a large window of opportunity for malicious exploits. The GitHub State of the Octoverse Report 2020[26] reported that vulnerabilities in Open Source Software (OSS) often go undetected for more than four years.

Some of the recent high-impact software security incidents have been caused by intentional tampering of the software supply process. Securing agile software development flow, built on the principles of Continuous Integration (CI) and Continuous Delivery/ Deployment (CD) is hard enough. Yet, it is even harder to assess the use of industry best practices for OSS development processes. The 2020 FOSS Contributor Survey from the Linux Foundation and the Laboratory for Innovation Science at Harvard University found that OSS developers spend insufficient time on security. It is advisable to pay attention to managing such risks e.g. by investing in regular source code reviews and educating OSS developers to practice DevSecOps, Vulnerability and Incident Management processes, and by favoring software originating from sources that are known to follow best practices in their CI/CD flow.

Mitigating security and privacy threats under 5G

The subject of security and privacy continues to invoke a passioned response and high expectations from citizens and governments alike. At the same time, information security and business continuity are top concerns among enterprises which are embarking on a digital transformation journey. It's imperative, therefore, that IoT is secure from the start, protecting personal data, business-sensitive information, and critical infrastructure.

Regulators are expected to walk a fine line between protecting privacy, safeguarding national security, stimulating economic growth, and benefiting society as a whole. To succeed with 5G transformation, industries need to gather competence, understand new threats, and learn how to mitigate them.

Protecting 5G end users requires private sector as well as policy makers to take a holistic approach to network security and not only focus on individual technical parts or layers in isolation (see section 2.2). For example, interactions between user authentication, traffic encryption, mobility, overload situations, and network resilience aspects need to be considered together. It also means that security needs to be considered interdependently across standards, vendor products, architectures, network deployments and operations, in order to protect the end users. It is also important to understand relevant risks and how to address them appropriately by ensuring mitigation that reinforces the security of deployed networks and not just merely shifts the risk.

2. Conceptualizing security in telecom networks

2.1 What is a telecom network and how does it work?

Telecommunication networks consist of four main logical network parts: radio access network, core network, transport network, and interconnect network.

The radio access network (RAN) is an instance of access network, and a major part of modern telecommunications. There are many types of access networks, such as the 3GPP access networks: GSM/GPRS, UMTS, EUTRAN, NG-RAN (5G), and non-3GPP access networks: satellite, Wi-Fi or fixed (wired) access network.

The core network can provide a number of services to subscribers that are connected via the access network into the core, such as telephone calls and data connections. The transport network keeps the access network connected with the core, and the base stations within the radio access network connected with each other. The interconnect network connects different core networks with each other. The telecom network transfers voice and data across the globe with high quality and consistency. User devices such as mobile phones can stay connected regardless of time and place, which is all possible thanks to standardized signaling systems and interfaces.

2.1.1 The mobile network information assets

Core network functions and management systems are critical assets in a mobile network. Affecting the core network or management systems may compromise the confidentiality, availability, and integrity of the entire mobile network services. Radio access network is also a sensitive asset, as it handles user data and may be placed in critical locations. At the introduction of edge computing certain core network functions are expected to deploy closer to the access sites, which makes the access also critical.

Data is one of the most important assets in mobile networks, subscriber data being the most critical one in this category. Subscriber data comprises of communication data (voice, text, and data sessions) as well as subscriber-related information, such as identities, location, subscription profile, and connection metadata (e.g. Call Data Records or signaling traces). To protect subscriber privacy, this data needs to be protected at storage and at transport.

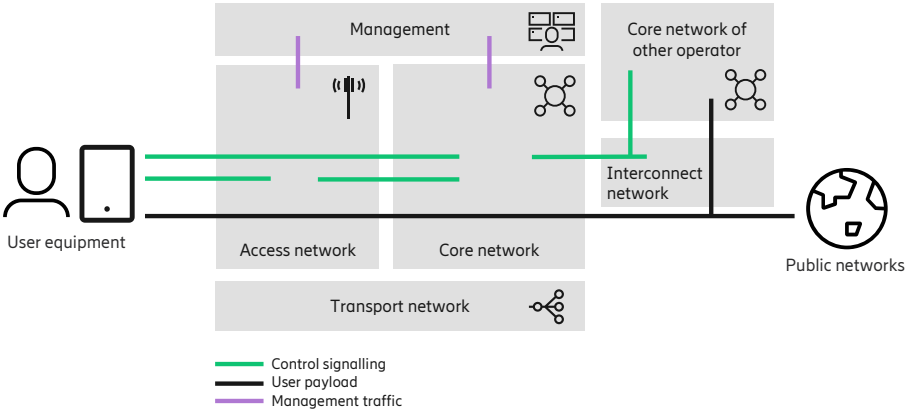


FIGURE 1: The telecommunication network – logical elements and logical planes

Apart from subscriber data, network management related information assets are required for proper operation of the mobile network. The management data comprises of infrastructure and service configuration data, network configuration data, security-related data, monitoring data, such as performance metrics, logs, and traces.

A prerequisite for data protection is to understand which data is critical, where does it reside and how it can be accessed. All data deemed critical must be protected over its entire lifecycle, including secure deletion. To ensure sustained protection, it is essential to enforce secure handling of encryption keys and the use of cryptographic algorithms and protocols of appropriate strength. In case of long-time data retention requirements, it may become necessary to re-enforce the protection with stronger cryptography.

For data-at-rest, the protection should include file system protection, encryption, integrity protection and strict access control. Additional controls are necessary for data-in-transit; traffic analysis to detect passing data to unexpected communication endpoints, transport layer encryption, and monitoring changes to router and firewall configurations. Data availability must be ensured by applying and rehearsing data back-up procedures.

Data-in-use protection may require selective actions on specific data items e.g. for privacy protection. Additional access controls can be applied, and some data elements may need to be removed, obfuscated, or anonymized. Appropriate operational security processes must be applied to re-identify data items when necessary for troubleshooting and incident investigation, Administrative access to critical data must be equipped with login and session security procedures that ensure non-repudiation; multi-factor authentication, strict command logging, and when necessary applying the "four-eyes" principle.

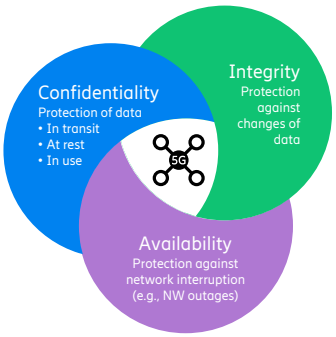


FIGURE 2. Protection of data

Each network part can be subdivided further into three so-called network planes, each of which carries a different class of traffic: signaling traffic, user payload traffic and management traffic. The signaling plane transports messages that are used to control user sessions, e.g. establishing a call or data session. The contents of a call or web page is referred to as user plane or user payload. The management plane includes management of monitoring, troubleshooting, configuration, and optimization of networks.
All planes are of interest for threat actors for varying reasons:

- Signaling[7] – the metadata which supports the networks is targeted to obtain information such as the geographical position of a subscriber. Modification of signaling traffic may be attempted to re-route calls or intercept SMS messages of a target for eavesdropping purposes or denying service. Today's security risks are far more developed and complex compared to previous generation technology. As such, signaling of previous generations, such as 2G, was developed with a reduced focus on security. This was owing, in part, to a high level of trust in signaling peers. Now we know better. Telecom signaling is regularly attacked and sometimes exploited on a daily basis. In current 5G 3GPP standardization, security is now taking a central role across all aspects.
- User payload traffic contains the actual data that is transferred for the user. Without appropriate security measures, the privacy of the user and the confidentiality of enterprise or government data would be at risk.
- The management layer is needed to ensure that the service provider's business performs optimally. The management plane is an attractive target for hackers to gain access to network resources, where they can manipulate and disturb network traffic and data. Mitigation of network management related risks and threats requires security policies and several security controls to be implemented, such as access control and security monitoring, in the right places (section 4).

Adopting this broader perspective ultimately leads us to encryption, something which is often mentioned in public debate. End-to-end encryption, although an integral tool, is still just one of the many tools needed to ensure the security of a system. Let's not forget also, the trustworthiness[8] of 5G does not only originate from a set of technical security features, but also from system design principles, implementation considerations and the day to day operations of networks.

2.2 Key security consideration in the standardization, development, deployment, and operations of telecom networks

Standardization[9] has played a vital role from the beginning of the emergence of global cellular networks such as GSM or 2G. In this process, operators and vendors agree about how networks around the globe will work together and how the networks and users can be protected against malicious actors. A key role of the standardization process is also to provide transparency on the agreed security solutions and specify the use of well-proven security algorithms, protocols, and architectures. Network vendors translate the agreed standards to functional network elements and systems. The design and development performed by the network vendor is a crucial part in making the end network product functional as well as secure.

In the deployment phase, the functionality and Network Functions as designed and developed by the vendors are integrated into the operators' target environment: the network. The network's architecture is designed and configured for a targeted security level as specified by the network operator. Activities to ensure network resilience includes e.g. setting security parameters and further hardening of the network. At the operational phase, operational processes which facilitate the network and deliver a targeted level of security is highly dependent on the deployment and operations of the network. One way of depicting these four interrelated processes is shown below in figure 3.

While the fundamental security features are specified in standardization, vendors enjoy a lot of room to maneuver throughout the development process, and so do operators throughout the deployment and operation processes.

Vendors implement common technologies differently.[10] Main features like interoperability and roaming are necessary, while non-common features (i.e. value adding features) differ from vendor to vendor. The quality and security of vendors' implementations vary and competition between vendors is an important driver for product level security.

A high level of product security assurance is vital for success in security. Security assurance, an important process in the vendor's software development process, usually contains a set of sub-processes on different levels to ensure that a product functions and performs as it is intended, and nothing else. Vulnerability assessment and penetration testing or risk assessment and privacy impact assessments are examples of such sub-processes. In addition, every piece of code needs to be reviewed and scanned for flaws and vulnerabilities. Security assurance is not limited to internal activities only.

Supply chain security controls often form a crucial part of a vendor's security activities. Similar standards of internal security need to be extended to suppliers of components and third-party software used in the end products and solutions. Most of the vulnerabilities exploited in live networks are publicly known vulnerabilities, often present in commonly used software components. Therefore, extra attention must be paid to monitor and respond to any vulnerabilities in any third-party components which are used.

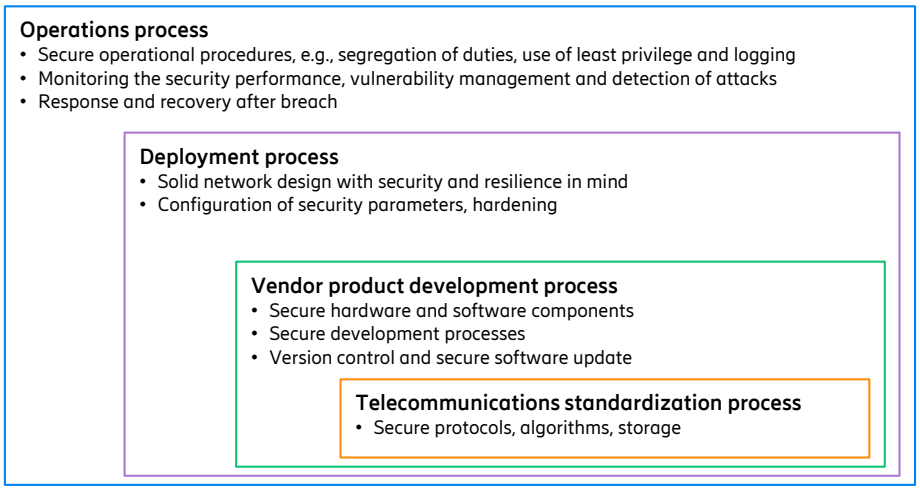


FIGURE 3: Key security consideration

2.3 What kind of general threats do telecom networks face?

The number of cyber security threats across the society and different industries has increased in the recent years. Telecommunication industry is no different and increasing number of attacks and attack attempts have affected the integrity, availability, and confidentiality of the infrastructure. Commonly these attacks are enabled by trivial security errors such as improper hardening, configuration, and usage of deprecated, vulnerable software versions. But in contrast, telecom networks also include components of bespoke specialized equipment and can only be targeted by malware which can be anything but trivial.

In large the technological barrier for successfully executing a cyber-attack doesn't exist anymore. Different types of malware and attack toolkits are sold as-a-service, complemented with options like trial periods, 24/7 user support, dedicated discussion forums and multi-language documentation. This development has contributed to a dramatic increase in the frequency of cybersecurity attacks, which have a low-risk and high-pay-off. Due to high degrees of digitization of industries and public services, the increased value of attack targets has also been aggravated by increased severity of impact that a cybersecurity attack can result in.

The threat actors behind cyber-attacks and their methods vary. When the motivation for the attack is money, the interest is high from criminals – deriving predominantly from malicious external actors, but also internal actors from within the network operators' or IT system organization (e.g. employees or subcontractors). For instance, having access to the billing and charging system of a telecom network allows insiders with malicious intent to commit fraud. Other typical attacker groups are hacktivists – politically motivated saboteurs who intend to disrupt service, deface websites, or steal sensitive information with the intent to cause financial damage or to send political messages. Another common class of attackers are insider threats such as disgruntled former employees or employees who seek to exploit their trusted position for personal gain.

While the same groups that target any other industry also attack telecom operators, telecom networks have some unique characteristics that makes them an interesting target for state sponsored actors and espionage. Telecommunication networks store and transfer location data and sensitive information like messages and voice conversations between high value targets, e.g., government officials, decision makers and high-ranking leaders. This information is of high interest to intelligence organizations from different parts of the world.

Industrial espionage has moved into the digital sphere as more and more of the valuable assets a company has are created, connected, stored, and shared, digitally. The goal is to gain access to a company's trade secrets, financial records, pricing information, intellectual property like new technology/innovations, and sensitive customer information. The uniting factor is the actor's objective of using information to swing the competitive situation in their favor. State sponsored actors have always had an interest in keeping an eye on what other states are doing. As social, economic, and political activities have increasingly moved to the digital space carried over public telecom networks, intelligence gathering operations have followed suit.

3. Critical infrastructure – increased value at stake in 5G

FIGURE 4: 5G use cases

5G will expand traditional relationships between consumers, business users and mobile network operators. The expansion will include new relationships in the form of digitized and automated business processes of enterprises, control, and operations of machinery of industry companies. Furthermore, cyber-physical interdependency between telecom networks and smart connectivity of other infrastructure providers (cities, power, utility, transport etc.) will be enabled by new ways to access the mobile network.

The 5G use cases for enhanced mobile broadband, fixed wireless access and cellular IoT (figure 4) are embodiments of new types of payloads carried over mobile networks. Although mobile broadband has been available and on the market for quite some time (mobile broadband was introduced in 3G), 5G is widely expected to introduce new qualitative and quantitative improvements, such as higher data rates, faster response times[11] (in the form of lower latency), more devices that can be simultaneously connect to a base station and higher bandwidth, across a wider area of geographical coverage. Furthermore, 5G also provides an increased level of security relative to 4G (see section 5).

The massive machine-type communication (which is a 3GPP term for IoT) will support tens of billions of power-constrained devices which typically transmit at irregular intervals, low volumes of data that are insensitive to delay.

Most industry stakeholders foresee a huge number of relatively simple devices that will need connectivity and create valuable data sets.

For applications which rely on 5G critical machine-type communication, they'll enjoy the benefit of ultra-reliable and low latency connectivity, where data volumes can be high and business critical. In this case, the communicating end-points are intelligent machines, vehicles, and robots with or without human interventions i.e. autonomous[12].

Industries and services which are expected to leverage such connectivity are: healthcare, manufacturing, transport, and consumer goods.

While IoT is a phenomenon that has already arrived and can be leveraged using both 4G and other non-3GPP access technologies, the machine type communication cases in 5G networks will empower IoT with network capabilities such as ultra-low latency which had not yet been available.

From a 3GPP network perspective, IoT means that mobile networks no longer connect only human identities in form of consumers, and business users, but also device identities. To achieve the necessary level of targeted security in mobile networks, the trustworthiness[13] devices must be considered which at minimum entails assuring the IoT device's identity[14], access control-essentially access privileges and confidentiality of associated data[15] (see figure 4).

While IoT device security is a critical component of an IoT security posture, it alone cannot ensure a secure IoT solution. Device classifications and security profiles are necessary to categorize each device type according to its use case, intended function, and classification of the data it processes and stores. IoT also introduces new security risks due to the number of devices, impact of an attack and lack of appropriate security controls. IoT solutions have unique security considerations because the devices are data-centric rather than human-centric. The IoT attack surface is across the entire IoT system, including the individual device profile, scale of devices, network interfaces, IoT application, IoT platform and shared resources in the cloud. A strong IoT security[24] posture takes zero trust and defense-in-depth approaches by placing security controls across the IoT system at multiple layers, protecting the end-to-end system and data to minimize risk.

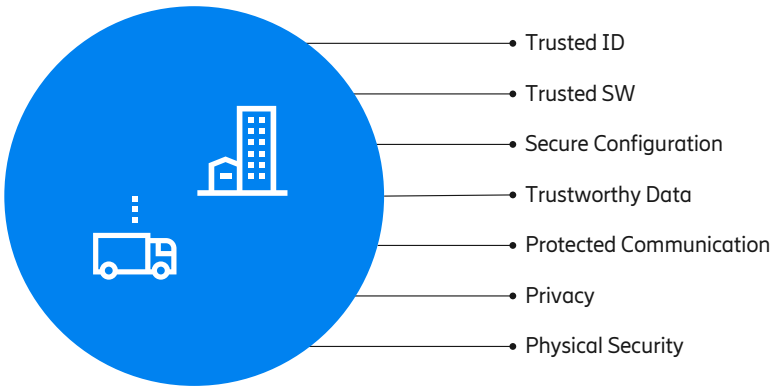


FIGURE 5: IoT device security aspects

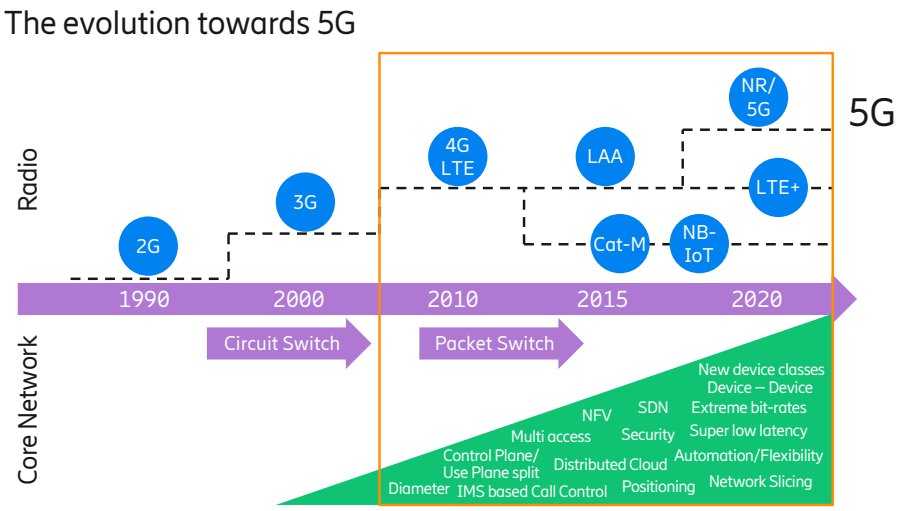


FIGURE 6: Evolution toward 5G and key technology trends

From the 5G network point of view, trust in IoT is based on trustworthiness of the device's hardware, software, configuration etc. Hence, trustworthiness is cumulative and will be defined by how well network operators and those who manage IoT devices govern the following:

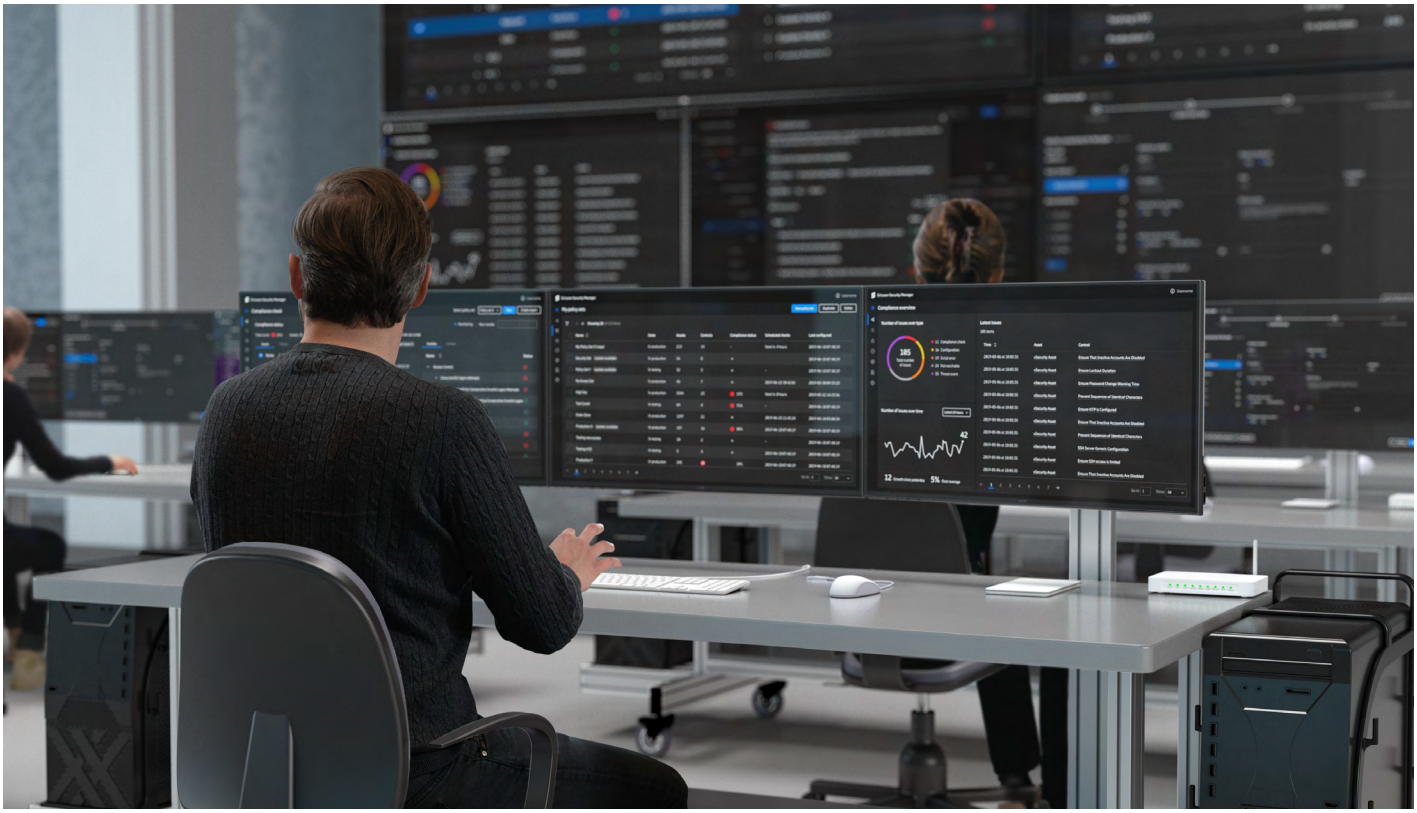
- identities and data,
- security and privacy,
- actor compliance with agreed security policies end-to-end.

Trustworthiness also depends on the right combination of trust enablers. For example, the hardware-based trust does not help if the application on top of the hardware does not make use of it. Ultimately, fully trusted application does not help, if the communication, e.g. the network between the applications cannot be trusted.

3.1 Key technology trends shaping the evolution of telecommunication networks

The 5G system may only appear as a faster and more versatile radio technology but it is much more. 5G is the first generation that was designed with virtualization and cloud-based technology in mind. The 5G system is not static for any specific access type or radio technology. For example, new services provided by the 5G core network are also available via 4G radio, Wi-Fi or fixed access depending on the network configuration. Evolution towards the 5G system started in the mid-2000s when the focus in telecom networks was shifted from circuit switched telephony services to packet switched networks and mobile broadband (figure 6).

With cloud-based technologies, software execution can now be disconnected from specific physical hardware (removing the need for boxed, e.g. hardware dependent functions). This is made possible thanks to Software Defined Networking (SDN) and Network Function Virtualization (NFV). SDN offers flexibility how to configure the routing paths between dynamically configured virtualized network functions. The introduction of AI and increasingly powerful computers, together with cloud technologies, will become a key driver of automation technologies. Consequently, the dominant tendency in these technology trends is already resulting in telecom networks becoming more and more software driven.



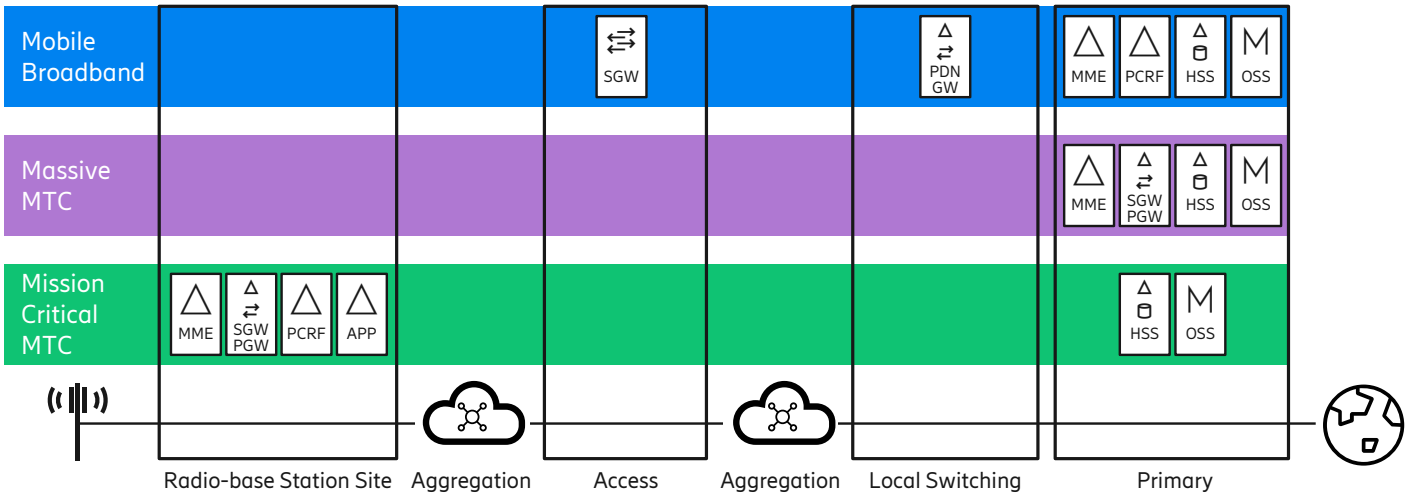


FIGURE 7: Network slicing

Distributed cloud computing makes it possible to create partitioning for better resilience and latency. From a security perspective, the distributed cloud may introduce new attack vectors against the 5G network if security is not built-in. On the other hand, distributed cloud may be seen as an opportunity, because of the possibility to place security functionality and mitigation mechanisms close to the attack source and thereby isolating the scope of the attack to local area.

The trend of connectivity, machine learning and other forms of AI is becoming more and more integrated across applications. Furthermore, market movement toward automation and autonomously controlled devices and vehicles is already beginning to take place at scale. Consequently, when these movements intersect, intelligent and autonomous devices will be an integral part of industry and society.

How all these technologies are built, integrated and controlled will become a major trust management issue for the future. Particularly for usage in critical infrastructures and to ensure privacy is protected. Network slicing[16] (see figure 7) is about separating different types of user traffic and creating dedicated core networks ad-hoc to facilitate a whole range of different 5G use cases (see figure 3). Network slicing enables the creation of device type, industry sector or even customer-specific subnetworks. The network slice control mechanism needs to provide appropriate slice management, configuration of access control, and secure isolation while still authorizing the shared resources. Each slice may have its own security policy that defines the security controls applicable for its specific threat landscape.

Network slices designed for critical services may also use the shared resources but require careful isolation. Critical services require high reliability, resiliency, safety, security and, often also, privacy. The security of critical services must ensure that communication parties and the connections remain protected. This requires comprehensive security approach including automated asset management and verification of security policy compliance. Open RAN's[23] goals of open, interoperable, virtual, automated, cloud-based networks can be achieved with vRAN, Cloud RAN, or O-RAN. While O-RAN's modified architecture introduces additional security risks due to its expanded threat surface, vRAN, Cloud RAN, and O-RAN share common security risks inherent to cloud-based deployments, for which the operator, cloud provider, and system integrator must establish a multi-party relationship to agree upon roles and responsibilities for implementing security controls. A zero-trust architecture, existing security tools, and industry best security practices for development, deployment and lifecycle management should be used to secure Cloud RAN.

4. 3GPP standardization of the 5G system

4.1 What is a 3GPP standardized 5G system?

The main service which the 5G system provides today's users is mobile (wireless) connectivity of a device to a network, often for Internet connectivity. This is also why the first 5G system use cases e.g. enhanced mobile broadband and fixed wireless access, are being deployed to offer users a better experience of the Internet.

3GPP does not typically standardize application services (such as Internet applications) since they are considered to be out of scope of 3GPP's connectivity focus. There are however a few exceptions: telecom networks have traditionally provided the possibility for two devices to connect to each other with the support of the network (e.g., to set up voice calls). In 4G networks, voice calls are set-up using voice over LTE (VoLTE) service on top of the connectivity service. VoLTE uses the IP Multimedia Subsystem (IMS) also standardized in 3GPP, similar voice service is also being developed for 5G.[17] Furthermore, 3GPP standardizes the security to support these services.

3GPP standards also cover some aspects of machine type communications and IoT. In 4G, the focus was to provide the cellular IoT devices with connectivity. In the 5G work 3GPP is investing a lot of effort to standardize support for industrial IoT and connected vehicle use cases. Consequently, the 3GPP standards cover efficient means to provide these IoT devices with an IP point of presence or even a complete 5G network within a factory to provide wireless connectivity to robots on the factory floor. To ease integration of 5G network within the factory, the 3GPP standard allows the re-use of existing authentication infrastructure of the factory. Security issues related to the actual application on top of connectivity are typically considered out of scope and need to be taken care of over the top. For example, 3GPP's 5G system can provide a temperature controller in a refrigerated goods wagon of a train with IP connectivity, but seen from the general 5G view, the authentication of the management traffic to the controller must be addressed over the top, since the IP address may be accessible via the Internet, so anyone could send messages to the controller. End to end security of the application layer is, however, not totally out of scope from 3GPP perspective. 3GPP has also defined features which enable to bootstrap keys from the mobile infrastructure for applications.

Apart from the security assurance specifications (see section 4.3 below), 3GPP does not standardize how mobile network system functions are implemented and realized. The main purpose of the specifications is secure interoperability between the functions required to provide network connectivity. Consequently, there has been little about virtualization and cloud deployments in the 3GPP specifications. However, the introduction of Service Based Architecture (SBA) with 5G [ref x2] has also brought some mobile network specific virtualization security aspects to 3GPP. In general, those aspects are handled through interaction with other standards organizations, especially ETSI ISG NFV (European Telecommunications Standards Institute, Industry Specification Group, Network Functions Virtualization) and ONAP (Open Network Automation Platform). Some details are not standardized at all and are left for implementations and deployments.

4.2 Security functions provided by the 3GPP standard

This section contains an overview of some of the most important security services provided by 3GPP standard to safeguard the connectivity for users, and the service availability and charging by the operator of the network. 3GPP's 5G system standards provide security mechanisms, which are based on well-proven 4G security mechanisms, but also include new enhancements for e.g. encryption, authentication, and user privacy.

While 3GPP security mechanisms provide reliable links for non-malicious bad radio conditions (see below) they do not protect against all possible threats, for instance DDoS and radio jamming. Protecting against DDoS attacks and radio jamming is something that is currently left for implementation and deployment, e.g. to re-route traffic via other base stations if one is jammed or scaling mechanisms and selective dropping/throttling in case of DDoS. Therefore, the appropriate level of cyber-resilience in the 5G system and 5G in general needs to be understood and addressed in a much broader way (see section 5) – 5G standards or, for that matter, any other technical standards will only be part of a much bigger picture.

Mutual authentication: Since 3G the end users of the 3GPP systems are authenticated to support charging for network access, accountability (e.g., which user had which IP address and when), and Lawful Intercept. The network is also authenticated towards the end-users so that the end users know that they are connected to a legitimate network.

Confidentiality of user plane data – the actual traffic data that is being transmitted – is achieved by encryption of end-user data as it passes through the mobile network to prevent eavesdropping over the air or on wires. 5G also introduced the possibility to protect the integrity of user data over the air. Once the data leaves the 5G system and traverses the Internet, the 3GPP standard does not ensure confidentiality nor integrity.

3GPP standards ensure that open, globally approved, and well-scrutinized encryption and integrity protection algorithm choices are made. 3GPP here relies on the support of security algorithm expert group of ETSI (European Telecommunications Standards Institute), specifically ETSI SAGE (Security Algorithms Group of Experts). For IP layer and above, 3GPP relies on well-proven IETF security algorithms and protocols.

Privacy threats to end users are mitigated by mechanisms that protect user identifiers. Note that, similarly to confidentiality, even though the 5G system protects the privacy of the end user using an Internet application

over the 5G system, the 3GPP standards do not intend to, and cannot, mitigate all privacy threats outside the 5G system even though there may be privacy concerns for the application also in a more general 5G setting. These threats require additional efforts by Internet application providers. The 5G system protects the messages sent by, for example, a social media user while they traverse through the mobile RAN and 5G system core network. The social media service must itself ensure that the message is protected end-to-end, since it will traverse the Internet once it leaves the 5G system. It is of course also up to the social media service to ensure the privacy of the user data once it has reached their servers and is being stored and processed.

The 5G system provides reliability and robustness against non-malicious unavailability situations, i.e. errors that appear due to unusual but expected bad radio conditions and broken links

A false base station[18] in GSM could identify a subscriber via the IMSI (International Mobile Subscriber Identity).[19] The technique is called IMSI catching. In GSM an attacker could even eavesdrop on users' data. Later generation mobile networks, starting from 3G, prevent the eavesdropping attacks because the network is also authenticated to the user. However, IMSI catching attacks are still possible in 3G and 4G. In 5G standards, even IMSI catching attacks can be mitigated. This is through a technique where the user's long-term identifier is never transmitted over the radio interface in clear text. Further, 5G increases the frequency with which temporary user identifiers are updated, further improving privacy.

The 5G system supports different types of compartmentalization, e.g. functions that aim to isolate possible security breaches from escalating from one part of the network to another. For example, there is a degree of split between the radio access network and the core network functions. This means that, should a radio base station get compromised, the core network, which provides global functions and processes more sensitive data, is still secure. Other examples of compartmentalization are cryptographically separated keys used at mobility events, and network slicing. Isolation of network slices is an important aspect, but it is currently not in the scope of 3GPP standards and is provided through implementation and deployment, e.g., targeted for specific use cases (see section 3) and desired performance and derived economic benefit.

Finally, one of the key purposes of 3GPP standardization is to ensure interoperability of security mechanisms between 5G system functions. Implementation aspects of the 5G system are only standardized by 3GPP to a very limited degree. For example, whether certain functions are implemented in single physical servers (physically isolated and separated) or implemented as virtual machines (VMs) or containers in a cloud or virtualized environment (shared hardware) is up to implementation and operator deployment choices (economics). This means that there is no simple rule of thumb derived from 3GPP standards regarding the separation of RAN and Core functions but rather flexibility prevails, even in a single physical network different configuration for different 5G use cases are possible, resulting in several differently configured logical networks are running over one physical network. For functions implemented, 3GPP in cooperation with GSMA, develops security assurance specifications, which set requirements for some implementation aspects.[20,21]

4.3 Security assurance in 3GPP SCAS and GSMA: NESAS

Mobile networks form the backbone of the connected society and are even classified as critical infrastructure in some jurisdictions, making security assurance especially important. Early on, the telecom industry realized the need to ensure secure implementations in addition to the secure standardized system and protocols. Therefore, 3GPP and GSMA took the initiative to create a security assurance scheme called the Network Equipment Security Assurance Scheme (NESAS)[22], which is suitable to the telecom equipment lifecycle. Ericsson strongly and actively supports the initiative in both 3GPP and GSMA by feeding the strongest parts of our own Security Reliability Model[27] (SRM) into the scheme, ensuring the other parts are covered by the scheme, and aligning the two.

NESAS comprises two main components: security requirements and an auditing infrastructure. The security requirements are defined jointly by operators and vendors in 3GPP. These requirements are currently defined on a network function basis and collected in so-called SeCurity Assurance Specifications (SCAS). There is, for example, one specification defining security requirements for 5G base stations. Various types of requirements exist, including the use of functional security policies, such as minimum length of management passwords, but also qualitative requirements on hardening and vulnerability testing. The auditing infrastructure is governed by the GSMA, the global mobile operator organization. The GSMA appoints audit firms that perform the audits of vendors' development and product lifecycle processes. The GSMA may also publish information on vendors that pass the audit of the development and product lifecycle processes.

NESAS aims to meet the needs of many national and international cybersecurity regulations, such as the EU cybersecurity act certification framework. The move towards larger portions of products being software – as we can see with SBA and cloud-based implementations – also offers the possibility for faster update cycles if vulnerabilities are discovered.

5. Security architecture in 5G

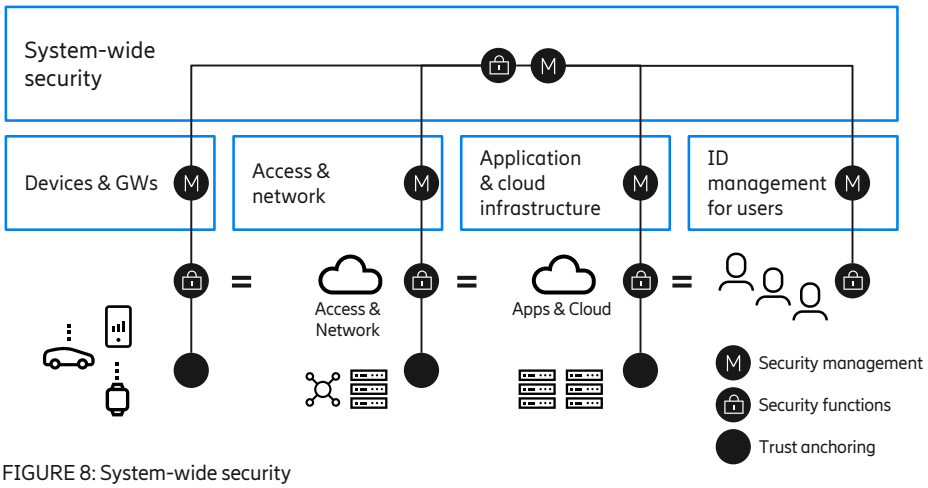


FIGURE 8: System-wide security

The 3GPP standardization section (4) focused on security mechanisms in scope for 3GPP, that being the functional elements and interfaces. Additional security considerations related to deployment scenarios of 5G system are covered in this section, including:

- System-wide security (horizontal security)
  - Network level
  - Slicing
  - Application level security
  - Confidentiality and integrity protection
  - Interconnect (SBA)
- 5G function element deployments (vertical security)
  - NFVi (virtualized or cloud native)
  - Appliance based functions
  - Distributed clouds and edge computing

5.1 System-wide security

As noted earlier, consumers and enterprises use existing (4/3/2G) cellular networks for mobile broadband (connectivity services), messaging service (e.g., SMS), and telephony services. Societal behavior and business services are evolving which raises the expectation on cellular networks to provide reliable and secure communication.

The aim of 5G is to become a reliable and trusted innovation platform for businesses and organizations to build and deliver new added-value services, but it is also considered an enabler for digitizing and modernizing critical national infrastructures such energy, transport etc. The latter raises the bar for 5G systems to provide greater availability and improved assurances of secure communication services. The horizontal, system-wide security approach spans across the network from the user device to the reference point where the operator terminates their services.

Horizontal security (see figure 8) is achieved by combining and coordinating a multitude of security controls across different domains in telecommunication networks, including radio access (e.g., radio unit, baseband units, antennas), transport networks (e.g., optical equipment, Ethernet bridges, IP/MPLS routers, SDN controller), packet core (e.g., MME, S-GW, PGW, HSS), network support services (e.g., DNS, DHCP), cloud infrastructure, and various management systems (e.g., network management, customer experience management, security management). Security across all these domains must be coordinated to provide the targeted availability of services and confidentiality and integrity of data sent, stored, and processed within the 5G system. Horizontal security will also protect the privacy of 5G users based on that data sent over the system is always confidentiality and integrity protected.

The previous section described the controls available in 3GPP nodes but let us now explore controls and design considerations in the transport and cloud domains in the 5G system.

Transport networks play an important role in the 5G system because they provide high-speed low-latency connectivity services between all 5G network functions. Consequently, the availability of transport networks is directly related to the availability of the 5G system and the services it provides. To ensure availability of transport services during node failure, cable or fiber breaks, or overload events transport networks can employ various technical solutions as well as considerations during network design, including:

- Geo-redundant paths that allow traffic to be re-routed in case of a path failure.
- Link redundancy solutions for fast failover in case of port or link failure.
- Path redundancy mechanisms that re-routes traffic flows due to path failure or overload conditions.
- High-availability configuration of critical network nodes to handle node failure.
- Use of traffic segmentation mechanisms (e.g., VLAN and MPLS) to logically separate traffic between different domains.
- Quality of Service enforcement using traffic queuing mechanisms, rate limiting, and traffic policing for resource and congestion management.
- DDoS detection and mitigation solutions.
- Port-based authentication to verify authorized network devices are attached to the network.
- IPsec or MACsec to create authenticated and cryptographic secured tunnels for sending data between sites and network elements.

The Service Based Architecture (SBA) and splitting of functionality in the traditional radio baseband unit opens to deployments in cloud environments. This flexibility grants several opportunities to realize new value-added service offerings, but also bares new risks and attack vectors that must be controlled in order to uphold the operator's targeted security posture. Some activities and controls to increase the trustworthiness in cloud include, but are not limited to:

- Hardening of the Network Function Virtualization environment, e.g., host OS hardening, secure configuration of the hypervisor or container environment.
- Tenant separation such that tenants are unable to interfere, have unauthorized data access, or intercept network traffic from other tenants.
- Compliance monitoring of tenants to ensure they remain within defined security policy
- Generation of detailed audit trails to support incident response and restoration activities
- Workload life-cycle management to ensure secure onboarding of virtual network functions, verify the integrity of software during boot and the integrity of workloads in operations, and secure decommissioning of workloads.

A logical construct that is used to describe the segregation of network services with different performance and security properties is the network slice. Several of the abovementioned controls and design guidelines will be combined to realize different network slices. For instance, a mission critical application that requires high availability, priority access to resources, and isolation from other services may be realized using services with geographic path redundancy with fast failover, authenticated with confidentiality and integrity using IPsec, and processed by dedicated 5G core network functions deployed on committed server blades and network security functions deployed with policies tailored to the specific application requirements.

5.2 Security considerations for O-RAN[2,28]

Architectural changes in 5G have created the opportunity for network virtualization to maximize flexibility and reduce costs to meet use-case-specific requirements. Virtualization means that security needs to be handled in a new way. As the industry evolves towards RAN virtualization, with virtual RAN (vRAN) or Open-RAN (O-RAN), it is important that a risk-based approach is taken to adequately address security risk. vRAN leverages the 5G split-RAN architecture, interfaces, and security protection mechanisms standardized by 3GPP. Building upon the foundation set forth by 3GPP, O-RAN is standardized by the O-RAN Alliance with new functions and open, interoperable interfaces. With any nascent technology, including O-RAN, security cannot be an afterthought and should be built upon security-by-design using a risk-based approach with consideration of potential zero-day attacks. O-RAN security controls include the following recommendations:

- Protect expanded threat surface due to additional functions and interfaces, including signed software and secure boot in RU and DU and integrity protection and encryption of management plane and control plane on the open fronthaul interface between the RU and DU.
- Close security vulnerabilities including root of trust for ORAN platforms, components, and applications. This is being worked in the O-RAN Alliance[25]. Associated examples: Near-RT RIC and root of trust for 3PP xApps and protection mechanisms from xApps conflicts.
- Implement a trust chain to mitigate risk of decoupled functions.
- Ensure management interfaces are secured according to industry best practices using TLS and digital signing.
- In addition, there are security risks introduced to O-RAN that should be mitigated for any virtual environment, such as
- Practice a high level of due diligence to protect again exposure to public exploits from use of Open Source code.
- Implement defenses for 3PP hardware that could be at risk of physical attacks

5.3 Deployment/Vertical security

3GPP specifies network functions and how they interact, but it does not specify how network functions should be implemented in embedded systems or in virtual environments.

Traditionally, radio base station equipment and radio core nodes are developed on vendor designed hardware platforms. These platforms have been carefully designed to meet strict requirements on availability, mean time between failures (MTBF), performance, scalability, power consumption, and physical security properties.

For example, a radio baseband unit includes tamper resistant hardware to securely store sensitive secrets, support secure boot procedures that verify the integrity and origin of software that is loaded onto the hardware, and hardware accelerators to boost cryptographic performance. During the manufacturing of baseband units, the hardware is provisioned with vendor unique credentials, called Vendor Credentials, that are used to cryptographically authenticate the device vendor of origin. This credential is used to secure deployment and integration of baseband units into operator networks. These credentials are securely stored on hardware devices with an established Trusted Execution Environment (TEE) as specified by the Trusted Computing Group creating trust that can be carried into deployment that is rooted in the hardware.

In virtualized deployments, the situation is different since multiple vendors may be involved in providing different parts of the solution, such as the hardware infrastructure, the virtualization platform, and the applications execute the 3GPP network functions. Secure provisioning and storage of identifiers and credentials is integral to provide a secure deployment in virtualized deployments. Currently, the industry is working on establishing methods to achieve similar trust and security as in embedded systems. For instance, hardware platforms (data center servers) need to include hardware technologies such as trusted platform modules (TPM), hardware security modules (HSM), and secure enclaves in CPUs and these capabilities need to be utilized by the virtualization platform and exposed to and attested by applications running on those platforms.

Virtualization of 3GPP functions allows a flexible distribution of the functions across infrastructure across the network in ways that are not possible for hardware-based solutions. It is possible, for example, to deploy a network slice where both RAN and core network functions are deeper in the network towards the edge on a distributed cloud platform to serve local enterprise services or regional IoT applications. This requires that network orchestrators which deploy the applications, the distributed cloud platform on which the applications run, and the applications themselves can be hardened and enforce required security controls. This is necessary to meet the operator's wanted security posture, at the same time as fulfilling the security requirement for the network slice use case. This is achieved by solutions that coordinate service deployment and security configuration across all involved domains. After deployment, security monitoring is needed to verify that the wanted security state is maintained throughout the lifecycle of deployed services.

6. Ericsson's 5G product security

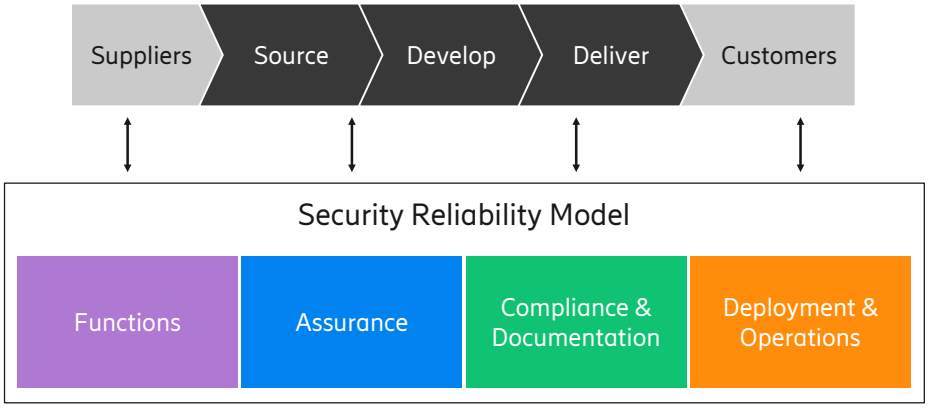


Figure 9: Ericsson Security Reliability Model

Ericsson's 5G network products build further on proven 4G platforms which today offer state-of-the-art security functions and advanced product security mechanisms.

Advance product security mechanisms, together with access management, logging, single software track and analytic tools constitute a solid foundation for implementing security policies and operating the network securely. The Ericsson Security Reliability Model (SRM) (section 6.2) framework specifically addresses operational needs by mandating hardening guidelines and security user guides for all Ericsson products. Additionally, Ericsson strategically offers products near security engineering services to assist operators in making network security assessments and configuring the network according to identified needs.

6.1 Key 5G security functionality

In addition to the improvements already described as part of the 5G standards, new technologies and deployment scenarios and use cases drive the need for applying state-of-the-art security technology. Ericsson is actively working in several areas to achieve this.

A fundamental challenge instantiating a Virtualized Network Function (VNF) is to securely provision it with roots of trust that enable it to become a trustworthy peer in the network that can protect the confidentiality and integrity of data both in transit and at rest. Here, Ericsson has developed solutions, founded in in-house research, and built directly into the 5G offering.

For Physical Network Functions (PNF), i.e. traditional HW/SW deployments, Ericsson's 5G RAN offering inherits the hardware rooted security for secure boot and signed software verification established already for 4G/LTE.

5G Core control plane has by standard got improved authentication and protection by mutual authenticated TLS. Ericsson SBA (Service Based Architecture) offers automated and trusted enrollment mechanisms with remote attestation and will use secure enclaves for protecting secrets. For protection of user equipment, a new firewall function is included in the UPF which monitors payload. Each container will have a unique certificate to ensure trust.

5G enables new use cases with more actors, applications and devices interacting with the infrastructure. In this new environment, efficient control of who may interact with whom, and who may do what and where, becomes a central security objective. To this end, Ericsson has already developed tools for efficient and correct policy management, policy distribution, policy verification, and policy enforcement that can enable functionality across tomorrow's networks.

Defense-in-depth is an important principle and what cannot be prevented must be detected, responded to, and recovered from. Telecom networks are uniquely instrumented to monitor performance in general. Ericsson is leveraging and augmenting these capabilities and together with modern analytics technology, drawing upon AI and Machine Learning, creating intrusion detection capabilities for our networks.

One specific concern that has received considerable attention is the ability to build a false base station through readily deployable technology and at relatively low cost. In this area, Ericsson has contributed toward standardizing functionality that will make it possible to efficiently detect the presence of rogue radio nodes in the network.

Perhaps one of the most important priorities for Ericsson is the relationship between usability and serviceability. Security functionality is of little value if it is not used or if it is used in ways that defeat its purpose. Here, Ericsson is working toward making security functionality unobtrusive and the default thing to do. Ericsson also engages with its customers to show and discuss how its products can be leveraged to contribute to reaching the customer's operational security objectives.

6.2 Ericsson's Security Reliability Model

Ericsson has a long history of systematically incorporating security and privacy considerations into all relevant aspects and phases of our product value flow. Our efforts in this area follow a well-established internal control framework known as the Security Reliability Model (SRM). The SRM enables a managed, risk-based approach to security and privacy implementation where requirements are tailored to the target environment and demands. This approach helps us meet stakeholders' expectations and cater for the rapid evolution of technology and the continuous changes in legislation globally.

The SRM defines Ericsson's approach to achieving our product security and privacy by design ambitions. Its purpose is to:

- set the product security and privacy ambition levels for our products and solutions
- specify the control framework that will enable Ericsson to fulfill the ambition levels
- outline how this control framework covers the product value flow, from the sourcing of components throughout the development activities to deployment and operations in customer networks
- demonstrate how we achieve compliance with relevant laws and regulations

The SRM divides security and privacy controls into four key areas: (1) Functions, (2) Assurance, (3) Compliance & Documentation and (4) Deployment & Operations. We apply the controls in these four areas continuously through well-defined activities across our product value flow, all the way from our suppliers to our customers.

SRM area #1: Functions
The SRM mandates that each product organization shall analyze, decide, and document the applicability and compliance to our Generic Product Requirements (GPRs) for security and privacy. To assist them in this work, it also defines a set of generic security and privacy functions for Ericsson products. Risk assessment and privacy impact assessment processes are used to identify and prioritize the applicable security and privacy functions from the set of GPRs.

SRM area #2: Assurance
Assurance refers to the activities conducted to ensure that the final product is secure when it is running in its target environment. They include risk assessments, privacy impact assessments, secure coding, vulnerability analysis and hardening. The SRM specifies the relevant assurance activities for each category in every phase of the product value flow: Source, Develop and Deliver. Depending on the characteristics of the product, the appropriate level of assurance activities – basic, advanced, or tailored – is set for each category.

SRM area #3: Compliance & Documentation
The Compliance & Documentation area covers all the information that demonstrates the security and privacy status at product release and in the customer documentation. It also defines applicable certificates and statements of conformance for external stakeholders, as well as providing necessary guidelines to maintain security and privacy in customer environments. Customer Product Information, security test reports and security standard conformance all play a key role in this area.

SRM area #4: Deployment and operations
The Deployment and operations area groups together the operational aspects of product security that arise in the product value flow, including security in system integration, guidance that operators require to operate their network in a secure way and customer support to resolve any incidents that arise.

Managing vulnerabilities
The work to avoid vulnerabilities includes product and feature risk assessments and secure design, secure coding principles and use of analysis tools, and supply chain security considerations.

The Ericsson process emphasizes the importance of risk assessments to identify needs for extra controls and to avoid functionality that could be abused by a malicious actor. Risk assessments serve to identify exposed parts of a product which require extra attention in coding and testing.

The use of secure coding principles contributes to overall code quality and robustness. Ericsson is committed to the idea that secure code is good code and that good code is secure code. Apart from mandating allocating time for programmers to learn about secure coding practices, Ericsson also provides the design teams with a wide selection of code analysis tools as part of the development environment and infrastructure.

Supply chain security considerations are prime concern for all industries and perhaps especially for telecom. For Ericsson, we believe it is business critical to address these concerns to the satisfaction of the customer. To this end, an internal program continuously works, with the support of senior management, to apply the standard risk management cycle of assessing risk, planning mitigations, deploying controls, and evaluating the results.

Detecting flaws
There is no way to entirely remove vulnerabilities in software they need to be managed and hence much of the total design effort goes into testing. However, testing for security vulnerabilities very much is about crafting input that lies outside what is expected and tested for normal operations, and that cause the system to misbehave in a way that can be exploited by an attacker.

Vulnerability analysis
A vulnerability analysis comprises the testing and verification of activities that are designed to identify weaknesses and vulnerabilities in the product or solution. The vulnerability analysis verifies the security characteristics and security configuration of the product/ solution and identifies new vulnerabilities through both black-box and white-box testing. Multiple tools and techniques are used, such as port scanning, vulnerability scanning, fuzzing and dynamic web application testing. Manual penetration testing is also performed. The verification of the security controls' functionality is done in the product's functional verification.

To design such testing, special competence is a prerequisite. Ericsson maintains Vulnerability Assessment teams that, with their knowledge, experience, and tools regularly prevent such flaws from graduating

to the release phase. Fuzzing is one technique that is used extensively to randomly introduce unexpected variations into protocol messages that are processed by a product. Where available, state-of-the-art commercial tools are used, but for more specialized interfaces, Ericsson works to develop in-house support for fuzzing and other tests methods.

When utilizing single software track handling principles for the whole software development and maintenance, vulnerability assessment before each software package release has two folded benefits. First, solutions or updated software patches for known and detected vulnerabilities can be mapped efficiently to the main, and the only, software track before releasing it commercially. Another benefit is that the vulnerability exploitation is fully avoided when the latest software version is deployed in the market.

Vulnerability watch

One enabler for building very complex systems is the utilization of high-performing third-party components and libraries. The reuse of proven code, both open source and commercially licensed, enables most software companies to concentrate on creating added value, rather than reinventing the wheel. Unfortunately, however, including third-party functionality comes at the price of third-party vulnerabilities. To address this challenge, Ericsson maintains a catalogue of third-party components used in our products. The PSIRT continuously monitors both public and subscription-based sources for alerts on discovered vulnerabilities in third-party software. This allows external vulnerability notifications to be mapped to Ericsson products. Where there is a match, an alert is sent internally to the affected product development organization that provides an analysis of how the reported fault impacts the Ericsson product in question. The alert analysis provides information of the severity and potential impact of the vulnerability.

Vulnerability remediation

If a product is affected by a vulnerability, a trouble ticket will be created. Appropriate remediation will be implemented and provided through standard support channels. Ericsson applies a single software track approach for new developments where new revisions of software contain both new features and corrections. Releases are pre-scheduled, but if urgently needed, unplanned emergency corrections can be made. Single software track maintenance means that both features and maintenance are developed in the same master software track and each consecutive software release contains both maintenance and feature development contents. In addition to obvious efficiency benefits, this single software track maintenance approach means also much more efficient preventive maintenance, better quality and better prevention of security vulnerabilities compared to the traditional maintenance with dedicated maintenance software tracks.

6.3 Observations from product security incident response

Ericsson's product security incident response team (PSIRT) is responsible for managing product security incidents across Ericsson's portfolio. As PSIRT experiences in security incident response regularly testify, the most common way to fail in security is to have shortcomings in the configurations of the network, elements of a network or poor network operational procedures. In such situations, breaches often go unnoticed due to lack of monitoring of log files and data flows. When an incident is noticed, the investigation becomes very difficult, if not impossible, due to lack of traceability. If many internal users have administrator permissions to the network or subsystems accountability maybe lost. Often also the log files are not protected and stored long enough, or backup restoration is not tested. The lack of basic security hygiene in operational procedures contributes greatly to increased risk of network security breach and exaggerate the damage in the event of a security breach. The same flaws may allow the attackers to hide their tracks effectively, resulting in increased difficulty addressing detection, attribution, and complete remediation.

Good network design in network deployment is needed to limit options to laterally extend attacks. Breach in security of one network component should not expose the rest of the network to the attacker. The principle of defense in depth explains how security controls must exist on every layer and every stage, necessary as no layer can be trusted fully i.e. there is no such thing as a 'secure internal network'. Solid operational procedures will include segregation of duties of network administrators and provide traceability back to every change and action done in the system. No one individual should have unaccountability in making significant changes to the system alone.

It is widely understood that prevention alone is not enough. Resources need to be assigned to active detection of attacks and respond in a time sensitive manner during and after an attack. Effective detection, response and mitigation combined with exercised incident response processes is essential to avoid security breaches.

Glossary

Cloud RAN
A virtualized RAN that is designed to be cloud native, built in a future proof architecture and incorporating key elements such as microservices, CI/CD and containerization

vRAN
5G becoming software-defined and programmable, generating additional RAN architecture flexibility, platform harmonization and simplification

Open RAN
Industry term for open radio access network architecture. A RAN with open interoperable interfaces, RAN virtualization, and big data and AI-enabled RAN

O-RAN
Refers to the O-RAN Alliance

5G
The fifth generation of wireless telecommunication technologies that targets high data rate, reduced latency, energy saving, and massive device connectivity with theoretical speeds up to 20Gpbs.

4G
Fourth generation mobile network that delivers downlink speeds of 1Gbps when stationary and 100Mbps when mobile.

3G
Third generation mobile telecommunications network that supports between 200Kps to 3Mbps data transmission speeds.

2G
Second generation mobile telecommunications network that supports up to 250Kpbs data transmission speeds (GSM and GPRS are typical examples).

3GPP
The 3rd Generation Partnership Project, a collaboration between groups of telecommunications standards associations.

Artificial Intelligence (AI)
The ability of a digital system to perform tasks commonly associated with intelligent beings.

Authenticate
The process of determining whether someone or something is, who or what it declares itself to be.

Active detection
The process of proactively identifying the occurrence of a breach.

Baseband unit
A subsystem in a telecommunications device that processes baseband radio signals.

Botnets
A network of computing devices infected with malicious software and controlled as a group without the owners' knowledge.

Breach
A security incident where the confidently, integrity or availability of a system has occurred.

Continuous integration (CI)
The practice of merging all developer working copies to a shared repository several times a day.

Compartmentalization
Functions that aim to isolate possible security breaches from escalating from one part of the network to another.

Core
The "backbone" network which provides the interconnect between other networks and systems to exchange information such as calls and data, including the special purposes servers and databases.

Distributed Denial of Service Attack (DDoS)
A denial of service (DoS) attack is a malicious attempt to make a server or a network resource unavailable.

Distributed Cloud
Interconnecting data and applications served from different locations.

Edge computing
Computation and processing of data is performed on distributed device nodes as opposed to primarily taking place in a centralized cloud environment.

Encryption
The process of converting information or data (plaintext) into encoded format (ciphertext) to prevent unauthorized access.

European Telecommunications Standards Institute (ETSI)
A non-profit organization that establishes telecommunications standards for Europe.

Evolved Universal Mobile Telecommunications System Terrestrial Radio Access Network (EUTRAN)
The air interface in an LTE cellular network.

Free Open Source Software (FOSS)
Software that can be accessed, used, modified, and shared by anyone. Open source is often distributed under licenses that comply with the definition of "Open Source" provided by the open source initiative and/or that meet the definition of "Free Software" provided by the Free Software Foundation.

Functional element
A manageable logical entity uniting one or more physical devices.

Hardening
Increasing product security by reducing its attack surface. Hardening ensures that the product is configured in a manner that minimizes the risk of unauthorized access and system compromise.

Hypervisor or container environment
The separation a computer's operating system and applications from the underlying physical hardware.

Internet Engineering Task Force (IETF)
The body that defines standard Internet operating protocols.

IMS
IP Multimedia Subsystem or IP Multimedia Core Network Subsystem enables the convergence of data, speech, and mobile network technology over an IP—based infrastructure.

Incident
An event that results in unauthorized access, loss, disclosure, modification, disruption, or destruction of data.

Interface
A shared boundary across which two or more separate components of a computer system exchange information.

Interoperability
A characteristic of a product or system, whose interfaces are completely understood, to work with other products or systems.

Internet of Things (IoT)
The interconnection via the Internet of computing devices embedded in everyday objects to enable them to send and receive data.

IP connectivity
A network or interface that supports Internet Protocol (IP) communications.

Internet Protocol security (IPsec)
A framework of open security standards that help ensure private, secure communications over Internet Protocol (IP) networks using cryptographic security services

MACsec
A security standard which defines connectionless data confidentiality and integrity on ethernet links.

Integrity protection algorithm
A software algorithm that is designed to maintain and assure the accuracy and completeness of data.

Latency
Delays in transmitting or processing data.

Layer
Level of abstraction in a network protocol stack.

Long Term Evolution (LTE)
A standard for 4G wireless broadband technology that offers increased network capacity and speed to mobile device users.

Lawful intercept
Facilities in telecommunications networks that allow law enforcement agencies with legal authorization to wiretap individual subscribers.

Logical network
A way of representing networks that have the same connectivity properties.

Massive machine type communication
Automatic data generation, exchange, processing, and actuation among intelligent machines on a large scale with the quality of transmitting low volume of non-delay sensitive information.

Mean time between failures
Predicted elapsed time between inherent failures of a system.

MEC
Mobile Edge Computing or Multi-Access Edge Computing (MEC), provides execution resources for applications with networking close to the end users, typically within or at the boundary of operator networks.

Metadata
Summarization information of data, for example the duration of a call or who was called.

Mobile Broadband
Wireless internet, often through a mobile telecommunications network.

Network slicing
Virtualization capability that allows multiple logical networks to run on top of a shared physical network infrastructure.

Domain Name System (DNS)
A method and infrastructure for converting alphabetic names into numeric IP addresses.

Dynamic Host Configuration Protocol (DCHP)
A protocol for assigning dynamic IP addresses to devices on a network.

Network Function Virtualization (NFV)
The visualization of network services that traditionally run on proprietary, dedicated hardware

Next Generation Radio Access Network (NG-RAN)
Infrastructure for 5G.

Port-based authentication
A mechanism to authenticate devices wishing to attach to local access network.

Radio Access Network (RAN)
Technology that connects individual devices to other parts of a network through radio connections.

Geographical redundancy
Replicates data between two geographically distant sites so that applications can switch from one site to another in the case of failure.

Path and link redundancy
An alternative channel of communication in the event of a failure.

Global System for Mobile communication (GSM)
Also known as 2G technology employed in second generation telecommunication networks.

Generic Product Requirements (GPR)
Ericsson's set of requirements that define the needed product functionalities, security and privacy related product documentation, and the needed evidence about security assurance activities.

Radio jamming
The deliberate jamming, blocking or interference with authorized wireless communications.

Radio unit
A remote radio transceiver that connects to an operator radio control panel via electrical or wireless interface.

Roaming
When a cellular customer makes and receives voice calls, send, and receive data when traveling outside the geographical coverage area of the home network.

Payload
The part of transmitted data that is the actual intended message.

Penetration testing
An authorized simulated attack on a computer system, performed to evaluate the security of the system.

Product Security Incident Response Team (PSIRT)
Ericsson unit that is responsible for actively and continuously monitoring new vulnerabilities and making sure they are fixed timely throughout Ericsson's portfolio.

Scaling mechanisms
Mechanism to increase or decrease capacity to meet the required demand at a given moment.

Secure coding
The practice of developing computer software in a way that guards against the accidental introduction of security vulnerabilities.

Selective dropping/throttling
A technique to discard or queue incoming traffic, often in response to network congestion.

Service Based Architecture
System architecture centered around services that can register themselves and subscribe to other services. Employed in 5G core networks.

Software-Defined Networking (SDN)
An architecture that aims to make networks agile and flexible that enables providers to respond quickly to changing business requirements.

Signaling (traffic)
The exchange of information between involved points in the network that sets up, controls, and terminates a call or data session.

Security Reliability Model (SRM)
Ericsson's methodology to achieve security and privacy ambition in products.

Quality of Service (QoS)
Technology that manages data traffic to reduce packet loss, latency, and jitter on the network.

Transport network
Connects the access network with the core or base stations with each other within the radio access network

Trusted Execution Environment
A secure area of a processor used to guarantee code and data loaded inside is protected with respect to confidentiality and integrity.

Topology
The arrangement of a network, including its nodes and connecting lines.

Trusted Platform Module (TPM)
A specialized chip used to carry out cryptographic operations like the storing of encryption keys to secure information which is usually used by the host system to authenticate hardware.

User plane data
The part of transmitted data that is the actual intended message.

Universal Mobile Telecommunication System (UTMS)
Also known as 3G.

Voice over LTE (VoLTE)
A technology that supports voice calls over a 4G telecommunications network.

Vendor credentials
Vendor unique information used to identify hardware such as radio base station so that it can be identified and trusted in a specific operator network and used for bootstrapping operator keys.

Virtualization
To create a virtual version of a device or resource, such as a server, storage device, network, or operating system.

Vulnerability
A weaknesses or gap in a system that can be exploited by threats to gain unauthorized access to an asset.

References

1 Ericsson.com. Security standards and their role in 5G. [online] Available at: https://www.ericsson.com/en/blog/2020/6/security-standards-role-in-5g

2 Ericsson.com. Security considerations of Open RAN. [online] Available at: https://www.ericsson.com/en/security/security-considerations-of-open-ran

3 Ericsson.com. What is 5G? – Ericsson. [online] Available at: https://www.ericsson.com/en/5g/what-is-5g

4 Ericsson.com. 5G systems - Enabling the transformation of industry and society – Ericsson White Paper. [online] Available at: https://www.ericsson.com/en/white-papers/5g-systems--enabling-the-transformation-of-industry-and-society

5 Ericsson.com. IoT connections outlook – Mobility Report June 2018. [online] Available at: https://www.ericsson.com/en/mobility-report/reports/june-2018/iot-connections-outlook

6 Ericsson.com. Open source software security. Ericsson blog post. [online] Available at: https://www.ericsson.com/en/blog/2021/1/open-source-security-software

7 Ericsson.com. (2018) Signaling security – Ericsson White Paper. [online] Available at: https://www.ericsson.com/en/white-papers/signaling-security

8 Ericsson.com. (2018). 5G security - enabling a trustworthy 5G system – Ericsson White Paper. [online] Available at: https://www.ericsson.com/en/white-papers/5g-security---enabling-a-trustworthy-5g-system

9 Ericsson.com. 5G standardization – Ericsson. [online] Available at: https://www.ericsson.com/en/tech-innovation/standardization/5g-standardization

10 Ericsson.com. Telecom Security Products and Solutions - Ericsson. [online] Available at: https://www.ericsson.com/en/security

11 Ericsson.com. 5G ultra-low latency propels jet engine manufacturing. [online] Available at: https://www.ericsson.com/en/networks/cases/5g-ultra-low-latency-propels-jet-engine-manufacturing

12 Ericsson.com. Bringing 5G business value to industry - Ericsson. [online] Available at: https://www.ericsson.com/en/trends-and-insights/consumerlab/consumer-insights/reports/5g-business-value-to-industry-blisk

13 Ericsson.com. (2017). Protecting the networked society - Ericsson White Paper. [online] Available at: https://www.ericsson.com/assets/local/publications/white-papers/wp-iot-security-february-2017.pdf

14 Smeets, B., Bergström, D. and Kristiansson, J. (2017). Secure brokering of digital identities. [online] Ericsson Research Blog. Available at: https://www.ericsson.com/research-blog/secure-brokering-digital-identities, and, Smeets, B., Englund, H., Sandgren, N. and Ståhl, P. (2017). Smart Contracts for Identities. [online] Ericsson Research Blog. Available at: https://www.ericsson.com/research-blog/smart-contracts-for-identities

15 Smeets, B. and Ståhl, P. (2017). Secure IoT identities. [online] Ericsson Research Blog. Available at: https://www.ericsson.com/research-blog/secure-iot-identities

16 Ericsson.com. Network Slicing – Ericsson. [online] Available at: https://www.ericsson.com/en/digital-services/trending/network-slicing

17 Keller, R. (2018). Voice in 5G system—architecture and EPS fallback. [online] Ericsson Future Digital Blog. Available at: https://cloudblog.ericsson.com/digital-services/voice-in-5g-system-architecture-and-eps-fallback

18 Norrman, K. and Kumar Nakarmi, P. (2018). Detecting false base stations in mobile networks. [online] Ericsson Research Blog. Available at: https://www.ericsson.com/research-blog/detecting-false-base-stations-mobile-networks

19 Norrman, K. and Kumar Nakarmi, P. (2017). Protecting 5G against IMSI catchers. [online] Ericsson Research Blog. Available at: https://www.ericsson.com/research-blog/protecting-5g-imsi-catchers/

20 Norrman, K., Teppo, P., Mononen, K. and Nilsson, M. (2014). Setting the standard: methodology counters security threats. [online] Ericsson Review. Available at: https://www.ericsson.com/assets/local/publications/ericsson-technology-review/docs/2014/er-security-assurance-3gpp.pdf

21 Ericsson.com Security in 5G RAN and core deployments. [online] Available at: www.ericsson.com/en/reports-and-papers/white-papers/security-in-5g-ran-and-core-deployments

22 GSMA.com. Network Equipment Security Assurance Scheme. [online] Available at: www.gsma.com/aboutus/workinggroups/working-groups/fraud-security-group/network-equipment-security-assurance-scheme

23 Ericsson.com Open RAN explained. [online] Available at: https://www.ericsson.com/en/openness-innovation/open-ran-explained

24 Ericsson.com Implementing Secure IoT Solutions: A systems-based approach to cybersecurity for cellular IoT, Ericsson, June 2021. [online] Available at:www.ericsson.com/en/internet-of-things/iot-security/implementing-secure-iot-solutions

25 o-ran.org Alliance O-RAN Minimum Viable Plan and Acceleration towards Commercialization. [online] Available at:www.o-ran.org/s/O-RAN-Minimum-Viable-Plan-and-Acceleration-towards-Commercialization-White-Paper-29-June-2021.pdf

26 octoverse.github.com The 2020 State of the OCTO-VERSE. [online] Available at: octoverse.github.com/

27 Ericsson.com. Ericsson Security Reliability Model. [online] Available at:www.ericsson.com/en/security/ericssons-security-reliability-model

28 Ericsson.com. (2021) Security Considerations of Cloud RAN. [online] Available at: www.ericsson.com/en/reports-and-papers/further-insights/cloud-ran-security

29 Ericsson.com AI in networks [online]. Available at: https://www.ericsson.com/en/ai

About Ericsson

Ericsson enables communications service providers to capture the full value of connectivity. The company's portfolio spans Networks, Digital Services, Managed Services, and Emerging Business and is designed to help our customers go digital, increase efficiency and find new revenue streams. Ericsson's investments in innovation have delivered the benefits of telephony and mobile broadband to billions of people around the world. The Ericsson stock is listed on Nasdaq Stockholm and on Nasdaq New York. www.ericsson.com

287-01 FBG 101 0955 Rev B