Ericsson White Paper
Uen 284 23-3269
April 2020

ERICSSON

Security in 5G RAN and Core deployments

The separation of RAN and core is critical to the evolution of 5G networks because gNBs (5G base stations) terminate the encryption of user data, except when it is encrypted externally and is beyond the control of an operator's 5G network. Currently, we do not have any standard rules or guidelines for the separation of RAN and core functions, and the 3GPP standards are largely flexible; however, the actual separation of RAN and core functions depends on the 5G use case(s) in question as well as the commercial strategy of the operator dictating the specific network deployment situation. In addition, technical developments and initiatives, such as distributed RAN, split RAN, O-RAN and CPRI/eCPRI consortiums, also fragment and distribute the deployment of RAN functions, entailing a number of security implications. This document aims to provide a brief technical description on the topic of RAN (Radio Access Network) and core network separation with regard to security in 5G. It considers both the 3GPP standards that specify approved technical industry agreements as well as the realizations of RAN and core functions in commercial 5G deployments, impacting the separation of the two functions. Security in the lower-layer split of the gNB (next generation gNB is also discussed.

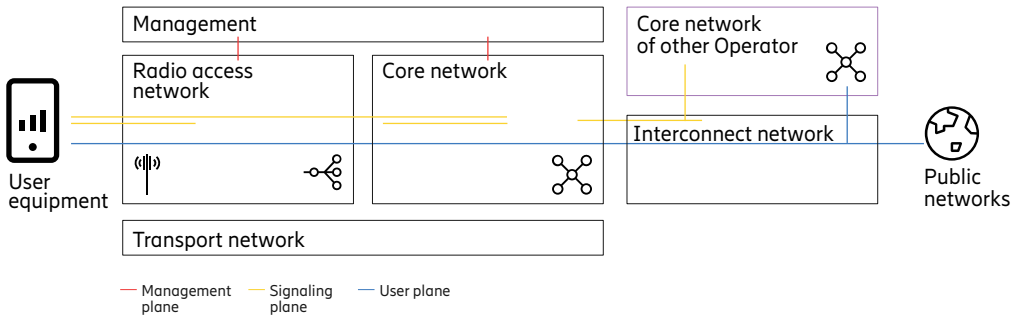
The main target audience for this document is the regulators, while the secondary audience is the decision-makers on the operator end.

Introduction

Telecommunication networks, as we know them, consist of four distinct parts: RAN, a core network, a transport network, and an interconnect network. These networks carry three distinct traffic types, commonly referred to as planes. The control plane carries the signaling traffic, the user plane the user data (which is the content of communications), and the management plane the administrative traffic. The administrative traffic contains configuration and control commands for RAN and core functions. Network security is critical to these planes, since all three of them are prone to unique and diverse types of threats. This paper focuses on the control and the user planes.

5G and the rapidly evolving associated technologies in the market put new security demands on telecommunication networks compared to previous mobile generations. 5G networks are critical to ensuring digitization and M2M (machine-to-machine) connectivity. We come across strategic terms like BOT (a software robotic device), IoT (Internet of Things), ML (machine learning), and so on that have the potential to drive the future of our industries. Naturally, then, there is much at stake, both in terms of value and risk.

With 5G, network security and privacy is improved, for example, by introducing IMSI (international mobile subscriber identity) encryption. All user data passing through 5G networks is confidentiality protected and integrity protected hop - by - hop.



The challenge

The security and trust level of deployed 5G networks is a topic of world-wide discussion.

Some papers, such as Ovum's The Facts on 5G argue that RAN is a largely insignificant part of a 5G network and cannot affect the confidentiality and integrity of 5G services.

Technically however, this is wrong, as the gNB is the termination point for encryption and integrity protection and, potentially, the user plane can be accessed in clear text (in case over the top end-to-end encryption is not used). Therefore, the user plane traffic is, in this case, accessible to anyone controlling the gNB or its implementation. In this paper, we would like to clarify that the technical aspect of security in RAN is as critical as the core network when it comes to confidentiality and integrity.

Regarding split RAN architecture, one could argue that some of the functions are not as critical when it comes to confidentiality and integrity, since the termination point for user data security (provided by the Packet Data Convergence Protocol, or PDCP) in the RAN is part of the CU (Central Unit). Since the RU (Radio Unit) and DU (Distributed Unit) can only access the user plane and the control plane encrypted, they pose no threat to user plane confidentiality . Still, we can see that without proper security functions in place, availability can be affected by the RU and DU.

Complicating the situation further is the fact that technical terminology is often unclear and is used in different ways by different parties, having the potential to mislead decision-makers and, thus, result in insecure systems. As a concrete example, some use the term "separation" to refer to the division between RAN and core functions in 3GPP standards, arguing, that there is a clear split between the two. Others, however, take this term to denote security. But separation as defined in the 3GPP standards assume that functions are securely deployed, properly implemented, and do not contain components with malicious intent. If that assumption fails, security does not necessarily follow.

Relationship between standards and deployments

3GPP has defined 5G architecture in TS 23.501 as illustrated in Figure 1 (below).

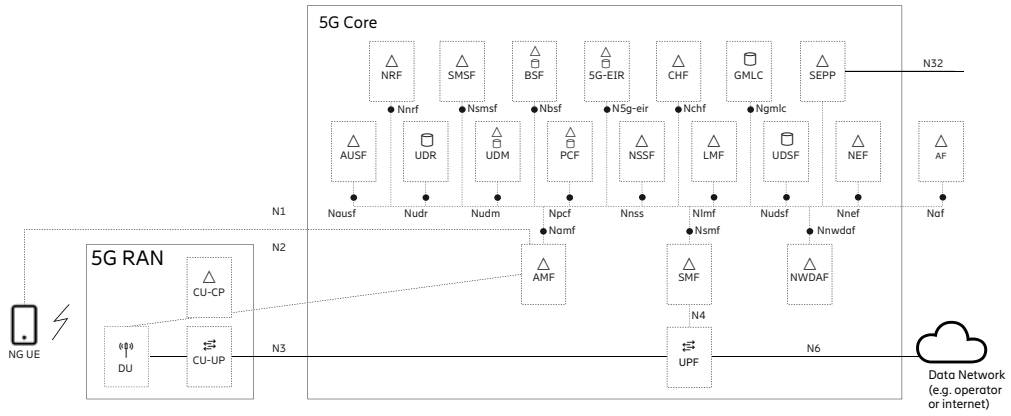


Figure 1. 3GPP 5G architecture

Within 3GPP standards, there is a clear functional separation between RAN and 5G core specified as the N1, N2 and N3 reference points. N1 uses N2 to transport its traffic within the network. N2 and N3 in Figure 1 shows the reference points within the network according to the functional architecture. In existing mobile networks, such as 3G and 4G, RAN and core have been deployed on geographically different sites. RAN has typically been deployed on several distributed sites to achieve optimized coverage and performance, whereas the core networks are deployed on a few regional and nationwide sites, supporting, in the process, many RAN sites. This conventional practice of deploying 3G and 4G networks is the reason why the separation of RAN and core is commonly presumed also for 5G.However, 5G and recent advances in implementation technology offer the possibility of flexible deployments, such as those with single sites, edge computing, cloud, and containers. These implementation and deployment choices are not defined by 3GPP standard. Rather, the degree of separation between RAN and core is established by commercial decisions. Consequently, service providers that seek to be commercially relevant for end users demanding low-latency services will not be limited by 3GPP standards. Instead, they will have the commercial incentives to deploy 5G networks in such a manner as to require only a very limited separation of RAN and core to achieve certain capabilities unique to 5G, such as low latency.

When discussing the separation of RAN and core in 5G, we also need to consider broader technological evolutions that have implications beyond 5G standards. In the case of 5G, the deployment of most network functions will be virtualized and cloud-based, allowing them to be deployed in different ways, depending on the use case. The mobile broadband use case, for example, may still rely on centralized core, but in cases related to IoT and manufacturing, a core network (or at least the user-plane function) closer to or co-located with RAN will be needed to support the low latency and high throughput required. Therefore, deployments with both RAN and core functionality on the same site prompt the question: is a separation between RAN and core really achievable in practice?

In addition to the co-location of RAN and core on a single site, edge computing also allows for the possibility of deploying other vital network applications on the same physical site, which further blurs the distinction between RAN, core, and IT cloud.

Security in RAN

5G can be deployed with either a 5G core network or by connecting 5G RAN to a 4G network. 4G is often referred to as LTE (Long Term Evolution). Security in RAN for these two different deployments is described below.

5G RAN with 5G core

Security in 5G networks is standardized (by 3GPP) in a hop-by-hop fashion, where user data is decrypted and encrypted in different functions within the network. User data is (in most cases) encrypted in transit (over the network) but processed in cleartext in many functions. The air interface is encrypted (and integrity protected) between the device and the gNB (5G base station). From the gNB over the backhaul network to the core network (normally via an edge router), the 3GPP defined NDS/IP security framework is used to protect the integrity and confidentiality of the user plane and control plane between the device, the gNB and the core network. The below figure illustrates how the different planes are protected in a 5G network.

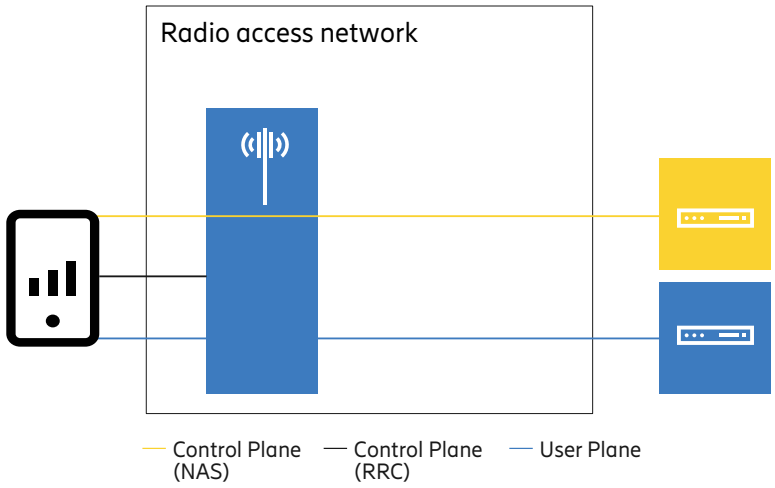


Figure 2. 5G RAN security

In a 5G (and 4G) network, NAS (Non-Access Stratum) signaling is encrypted between the device and the core network. Moreover, both the control plane (Radio Resource Control, signaling between the device and RAN regarding radio configuration) and the user plane are encrypted and integrity protected between the device and the gNB (or a base station called eNB in the 4G case), meaning that all user data is available unencrypted in the gNB(or eNB) . In many cases, user data can be encrypted at the application level, but this is not guaranteed by 3GPP 5G standards and is out of operator control.

Non-standalone 5G deployment

Initially, 5G radio will, in many cases, be deployed in a non-standalone fashion. In this situation, the gNB is connected not to the 5G core network but to an eNB in a 4G RAN , as shown in Figure 3. The eNB plays the role of a master base station, and the gNB plays the role of a serving base station. This setup is referred to as DC (dual connectivity). The serving gNB forwards the uplink user plane data to the master eNB, and user data is then decrypted in the RAN before being forwarded to the core network. This means that user data is available unencrypted in the eNB when this deployment is used.

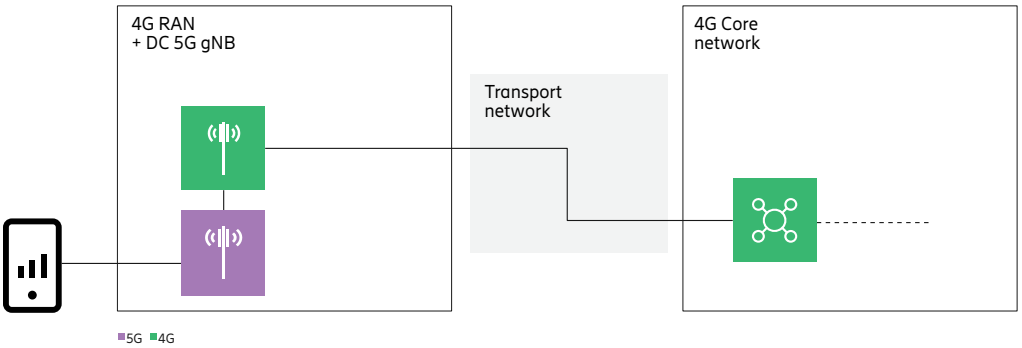


Figure 3. Non-standalone 5G RAN

Security in future RAN deployments

As discussed in Section 3, in 4G and earlier generations, base stations were designed to perform singular functions (normally implemented as separate physical units). In 5G, however, this is different. The ongoing development to separate the gNB in different functions is essentially aimed at deploying gNB functions in different ways. The 3GPP TS 38.401 specifies the possibility of a distributed gNB with a CU and DUs, as shown in Figure 4 (below).

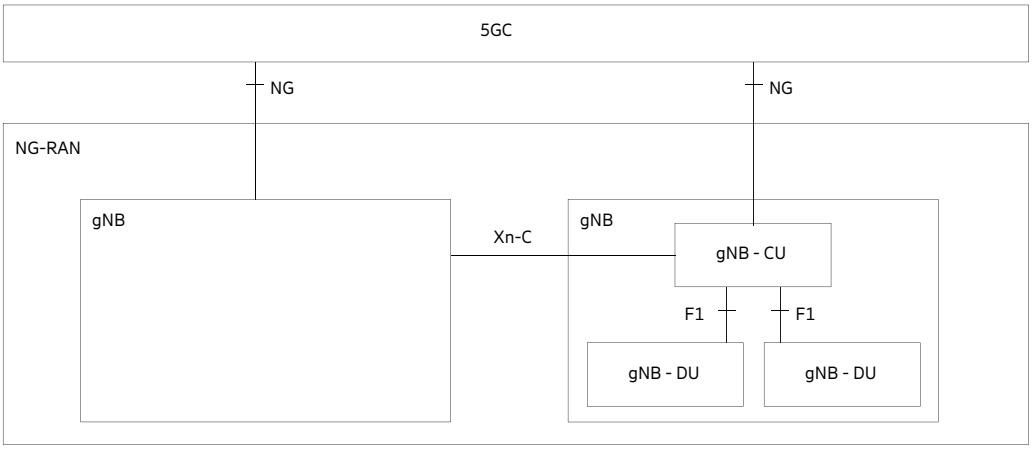


Figure 4.3GPP distributed gNB

The DU and CU are functions in the 3GPP-standardized 5G RAN. So, contrary to previous deployment conventions in 4G and earlier generations, in 5G, these RAN functions can be placed in different physical sites in an actual deployment of RAN, depending on the use case. This enables RAN function distribution over different physical sites and, subsequently, allows a breakout of RAN functions to support low-latency use cases as well as flexible implementations. Consequently, the split between RAN and core may be clear in standards, but it becomes unclear when viewed in actual deployments.

Other organizations than 3GPP, such as O-RAN aim to define implementation and deployment architectures, focusing on how we can further split RAN into even more granular building blocks than described above. Specifically, the work on the lower layer split (RAN functions) also introduces an RU in addition to the DU and CUs. This development allows even more ways of distributing RAN functionality, and therefore further blurs the distinction between RAN and core from a security perspective.

CPRI/eCPRI (common public radio interface/ enhanced CPRI) is another consortium specification for an interface and protocol between the RU and the baseband. It is currently not fully specified as a multi-vendor interface, meaning that some integration work is needed if the RU and DU are from different vendors.

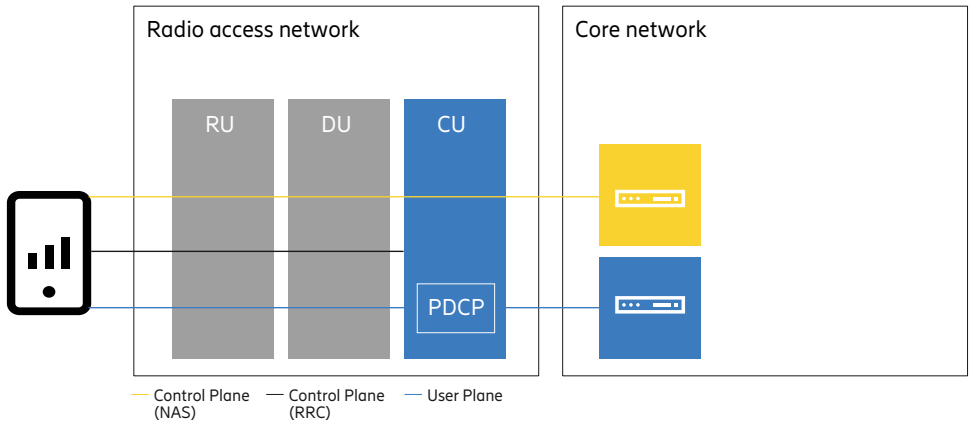


Figure 5. Security in split RAN

With the lower layer split, the termination point for encryption the CU function, which terminates PDCP on the network side. With this split, the RU and the DU are not able to access (that is, decrypt) the user plane and control plane, meaning that the RU and the DU are not as critical as the CU when it comes to the integrity and confidentiality of user data or the signaling. Still, both the RU and the DU can affect the availability of mobile network access.

Conclusion

Telecommunication networks are evolving every day. Advanced technologies such as 5G, IoT, and virtualization services may also affect the security of the network. 3GPP 5G standards allow physical and virtual overlap between RAN and core networks in deployed networks.

The separation of RAN and core is critical to the evolution of 5G networks, and may pose hurdles in securing the low latency use cases that have been important drivers for 5G development. Currently, we do not have any standard rules or guidelines for the separation of RAN and core functions; 3GPP standards allow for flexibility. The degree of RAN/core separation in a specific network deployment situation is not uniquely determined by 3GPP standards. The actual separation of RAN and core functions depends on the 5G use cases and the commercial strategy of the operator dictating the specific network deployment situation. To achieve the lowest latency, co-location of RAN and core may even be necessary.

When discussing the degree of separation between RAN and core in 5G, we also need to consider the technology evolution that is taking place around us and that goes beyond 5G standards. For example, network function virtualization and cloud-based deployments greatly blur the distinction between RAN and core.

RAN and core are both critical components of 5G networks because gNBs (5G base stations) terminate the encryption of user data, except when it is encrypted externally and is beyond the control of an operator's 5G network. As a result of this, gNBs have full access to all data to and from devices in cleartext. Moreover, technical developments and initiatives, such as distributed RAN, split RAN, O-RAN and CPRI/eCPRI consortiums, further fragment and distribute the deployment of RAN functions, with serious security implications. For instance, all the options make it unclear how functions will be distributed and co-located in the long run. There is a risk that market demands will drive the most cost-effective function distribution, not necessarily the most secure one.

Authors

Patrik Teppo
Expert - Security

Patrick joined Ericsson in 1995 and is currently working as a security expert in the CTO office. He is the driver for the security architecture implemented on our products and solutions. He holds a B.Sc. in software engineering from Blekinge Institute of Technology, Sweden.
Email: patrik.teppo@ericsson.com

Karl Norrman
Master Researcher - Security

Karl holds an M.Sc. in computer science from department of mathematics at Stockholm University and has been with Ericsson Security Research since 2001. He was actively involved in the LTE security standardization and was Ericsson's security coordinator in 3GPP. He currently works as a master researcher focusing on 5G security and is part-time pursuing a PhD in theoretical computer science at KTH Royal Institute of Technology, focusing on automated cryptographic protocol verification.
Email: karl.norrman@ericsson.com

References

- 3GPP 5G security
- 3GPP 4G security (Specifies one option of security in non-standalone 5G)
- The Facts on 5G
- 3GPP 5G architecture
- 3GPP 5G RAN architecture

Glossary

- 5G: fifth generation wireless targets high data rate, reduced latency, energy saving, and massive device connectivity.
- 3GPP: the 3rd Generation Partnership Project, a collaboration between groups of telecommunications standards associations.
- Baseband unit: a subsystem in a telecommunications device that processes baseband radio signals.
- Core: The "backbone" network which interconnects other networks and systems to exchange information such as calls and data, including special purposes servers and databases.
- CPRI: Common Public Radio Interface.
- eCPRI: published after CPRI, the eCPRI standard is used for LTE Advanced and 5G networks.
- Encryption: the process of converting information or data (plaintext) into encoded format (ciphertext) to prevent unauthorized access.
- 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.
- NDS/IP: Network Domain Security/IP is a framework for securing data in transmission between sites in a mobile network.
- Latency: delays in transmitting or processing data.
- Layer: a 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.
- Radio access network (RAN): technology that connects individual devices to other parts of a network through radio connections.
- Radio unit: a remote radio transceiver that connects to an operator radio control panel via an electrical or wireless interface.
- Service based architecture: system architecture centered around services that can register themselves and subscribe to other services. Employed in 5G core networks.
- Transport network: connects the access network and the core or base stations with each other within the radio access network.
- User data: the part of transmitted data that is the actual intended message to or from the user's device.
- Virtualization: to create a virtual version of a device or resource, such as a server, storage device, network or operating system.