A novel approach to radio protocols design for 6G

White paper

While discussing future use cases and deployment scenarios, it is very important to know where 6G is heading and the requirements 6G needs to fulfil. It is also crucial to consider early on the new requirements 6G will mean for implementers. This white paper reflects on how radio protocols have evolved since 2G, discusses the recent challenges 5G has faced, and offers a paradigm shift for the design framework of 6G radio protocols. The new framework proposes to rely on a dual stack approach, with a first stack hosting control plane functions and optimizations for low bitrate services, and a second stack optimized for parallel processing on radio processing units.

Contents

Introduction 3

Six goals for 6G protocol stack design 4

Protocol design 5

Overall framework 5

RPU management 7

Control plane 8

User plane 8

Physical layer 9

Conclusion 10

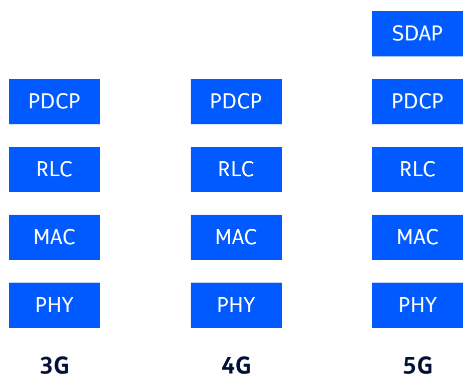
References 12

Introduction

Ever since the introduction of 2G packet data services (GPRS: General Packet Radio Service), it has become a habit to use the protocol stack of the earlier generation as a starting point for the design of the radio protocols of the next generation. This toolbox approach configures the one radio protocol stack to support a wide range of services — one stack to rule them all.

For instance, the long-term evolution (LTE) radio protocols were primarily designed for the provision of packet switching (PS) services through a flat architecture. They represented a major improvement over the previous generations, shedding the complexity inherent in the support of circuit switching (CS) services and a convoluted architecture. Many of the original principles of LTE have remained untouched since 3GPP Release 8. In the early days of 5G standardization, it was agreed to use the LTE radio protocols as the baseline for 5G and to enhance them to support very high data rates with low latency, dynamic spectrum usage and flexible quality of service (QoS) [1]. The one stack approach remained and, with a very large number of services to support for 5G from day one, the size and complexity of the stack increased dramatically [2, 3, 4].

Figure 1. 3GPP radio protocols



The main issue with the one stack approach is that while some optimizations can be considered as desirable for low bitrate services, the very same optimizations become irrelevant and even harmful when dealing with very high bitrate services. While it is perfectly acceptable to run some types of optimizations at low bitrate, running the same optimizations for gigabit services either becomes irrelevant (e.g., overhead optimizations), too costly (e.g., power consumption and hardware dimensioning), or both.

The ever-increasing demand for high bitrate services, increased security, and reduced power consumption [5] poses even further challenges to the one stack approach for the design of radio protocols. In this paper, we will present the Nokia vision of the viable principles for the design of the next generation of radio protocols that could overcome these challenges.

Six goals for 6G protocol stack design

Before explaining our vision of the design principles for 6G, it is important to set our goals. The figure below summarizes the six goals we have for the radio protocols of 6G: simple, scalable, smart, sustainable, secure, and stable.

Figure 2. Six goals for 6G

1. **Simple**: there are obvious benefits in keeping radio protocols simple. Simpler protocols are easier to implement, less costly to debug and faster to market. Another significant advantage, often overlooked, is that they consume less power. A valid analogy can be found in the history of microprocessor architectures. CISC (complex instruction set computer) processors were initially replaced by RISC (reduced instruction set computer) processors (except in Intel-based x86 computers), because they were simpler to program and easier to scale. ARM processors (Advanced RISC Machine) now dominate in mobile devices (both iOS and Android) because they save on power. As energy efficiency and sustainability are key drivers for 6G [6], the simpler architecture becomes the preferred alternative.
2. **Scalable**: the support of very high bitrate services highlights the presence of possible bottlenecks in the one toolbox approach. Processor architectures again provide an analogy. In the early 2000s, increasing the clock speed of a single core became a lot less efficient (both in terms of overall processing power and power consumption) than increasing the number of cores with slower clocks. Radio protocols encompassing the notion of radio processing units (RPUs) for parallel execution are expected to provide the same benefits.
3. **Smart**: the radio protocols should support AI/ML natively. The need to manage multiple RPUs also requires native support for increased intelligence in how the radio protocols operate [7].

4. **Sustainable**: the need to reduce power consumption to increase battery life and reduce environmental impact has never been so great. With the notion of RPUs, it would become possible to activate the cores on demand, thus low data rates would only require a single core for maximum power savings.
5. **Secure**: ciphering location has evolved over the years. In GSM, ciphering for CS was done after channel coding and interleaving but before modulation. With the introduction of GPRS and packet switching services, not everything could be ciphered. Header fields, for example, had to be decoded by several UEs. Ciphering thus moved up to medium access control (MAC) and radio link control (RLC) in 3G. In LTE, ciphering was first considered in the core network (CN), then moved back to radio access network (RAN) where it also stayed for 5G. The increased need for privacy and secure communications in a post-quantum world stresses the need for even stronger security and opens the door to discussing where security should be located [8].
6. **Stable**: with the possible exception of GSM, the high complexity of the first releases has always required subsequent releases to take steps to reduce the complexity, for example, machine type communications (MTC) and narrowband internet of things (NB-IoT) for LTE, and reduced capability (RedCap) for New Radio (NR). The radio protocols of 6G should be simple enough to avoid this, for instance, by allowing simple implementation with only one RPU from the initial release.

Protocol design

Having established the six goals, we will now describe the principles for the design of the next generation of radio protocols. We will first focus on providing radio protocols that are simple, scalable, sustainable, and stable. We believe that such a baseline can then easily support mechanisms to make the radio protocols secure and smart.

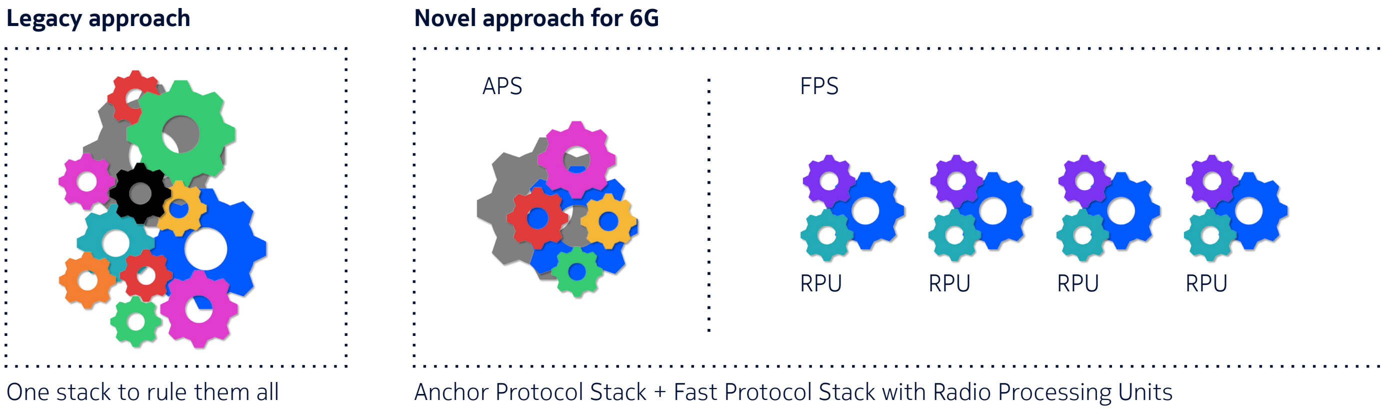
Overall framework

As described in the introduction, the optimizations which are required for low bitrate services become irrelevant and perhaps even harmful when dealing with very high bitrate services. The coverage of low bitrate services will remain crucial in 6G though; thus these optimizations will still be valuable. It is also important to acknowledge, however, that they may not always be needed, especially when they become harmful to very high bitrate services. Thus, instead of having one complex stack mixing all mechanisms and optimizations, we would like to suggest a two stack approach (see figure 3):

1. One radio protocol stack is designed for low bitrate services, coverage (e.g., bit-level optimizations) and reliability (e.g., RLC ARQ). Let us refer to this stack as the Anchor Protocol Stack (APS).
2. The second radio protocol stack is designed for high bitrate services, where the focus is on a processing-friendly and implementation-friendly design employing the concept of RPUs outlined earlier and leaving aside some optimizations that an implementation can only afford for low bitrate services. Let us refer to this stack as the Fast Protocol Stack (FPS).

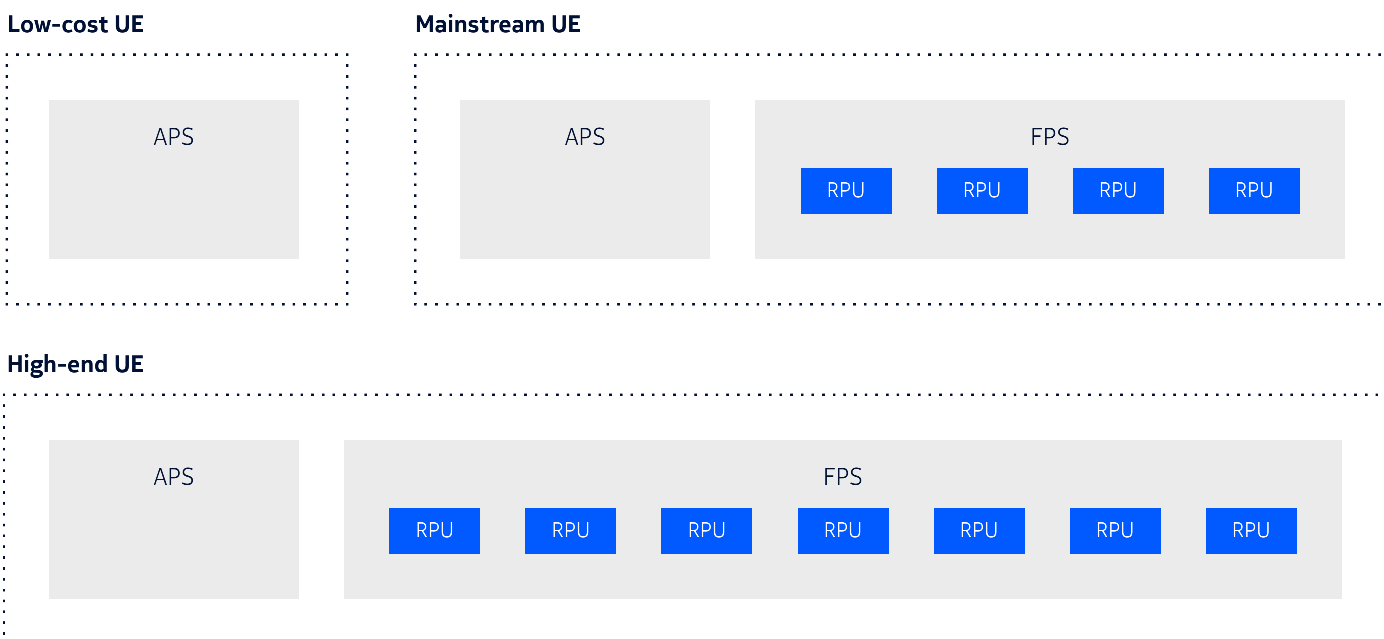
A simplified representation of this approach is depicted below where each cog represents a radio protocol mechanism (the larger the cog, the less suitable the mechanism is for high bit rate services). On the left side, the legacy approach with one radio protocol stack, and on the right side, the novel approach we suggest with two stacks and radio processing units.

Figure 3. A novel approach to 6G radio protocols



With such an approach, the complex mechanisms and optimizations that are fully justified for low bitrate services need not be used for very high bitrate services. A simple device may only implement the first stack (APS), possibly removing the need to introduce the equivalent of MTC, NB-IoT and RedCap. A more complex and capable device would implement both stacks. The higher the bitrates the device supports, the larger the number of RPUs the FPS would incorporate, as exemplified in figure 4 below with three types of UEs depicted.

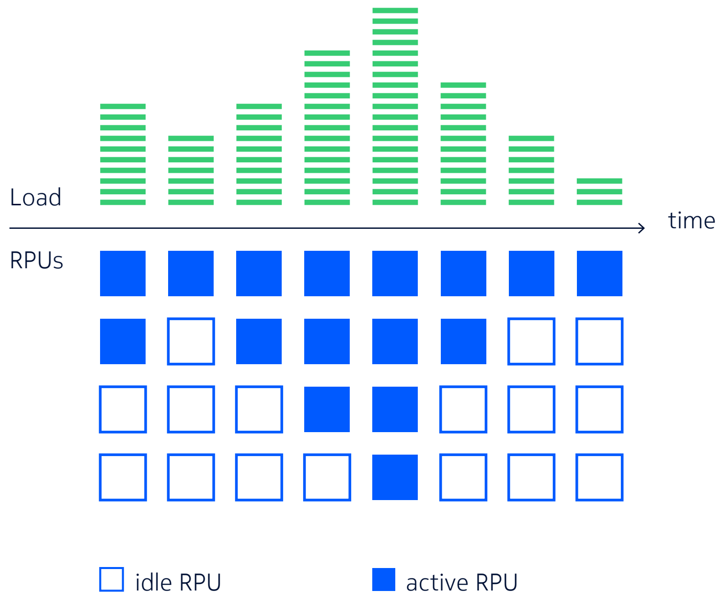
Figure 4. UE types for 6G



RPU management

The goal of efficient RPU management would be to maximise the power saving gains made possible by the RPU framework. The number of RPUs that are activated can be adjusted according to the instantaneous bitrate or load, as exemplified in figure 5 below, where a total of four RPUs are assumed to be available.

Figure 5. RPU management in 6G

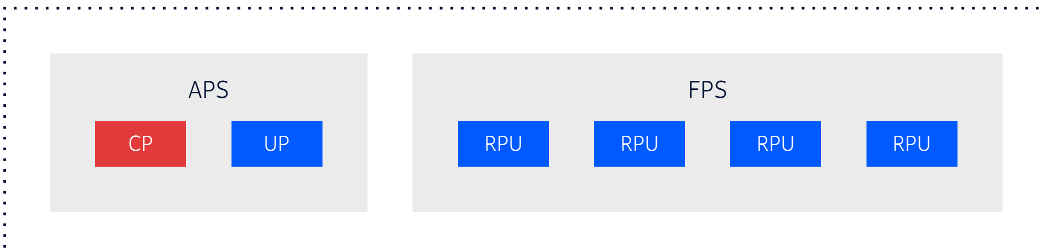


Ideally, such details should be left up to implementation. However, depending on whether the RPUs share a common memory and how they are activated, it is possible that some RPU management schemes would require specific mechanisms to be introduced in standards. For instance, if the RPUs operate on segregated memory resources, it is likely that each RPU would then host its own transmission and reception windows, thus impacting sequence numbers and status reports management. Conversely, RPUs operating on shared resources would allow common windows to be used, with no impact to sequence numbers or status reports.

Control plane

Being always present, the APS is a logical host for the control plane functions such as idle mode, connect mode and related configurations of the radio resource control (RRC). By containing all control plane (CP) functions within the APS, not only is the FPS free to focus on user plane transfer for a simplified design, but it need not be active when the bitrate requirements are low.

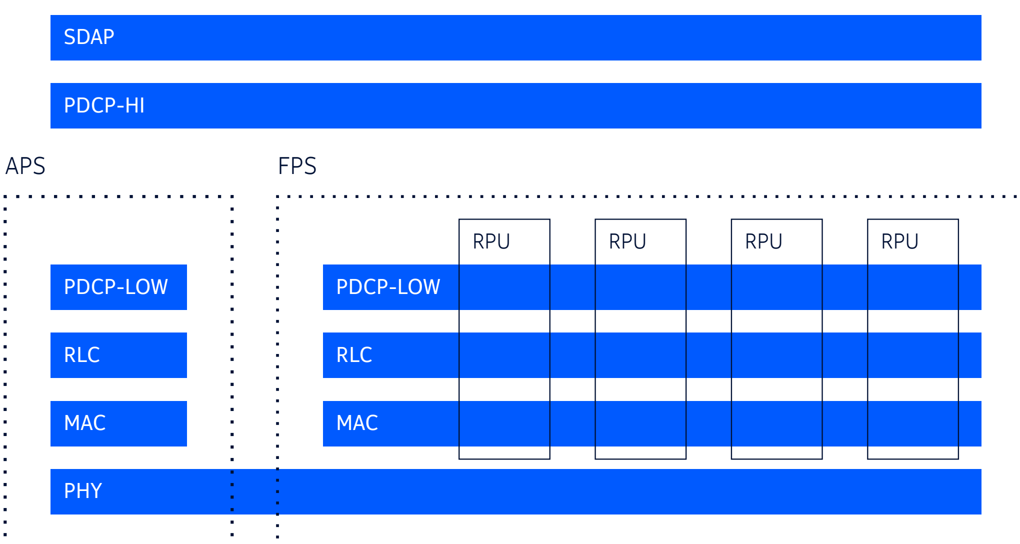
Figure 6. Control plane location in 6G



User plane

To allow parallel processing within the radio protocols of the FPS, each RPU needs to host sublayers of layer 2 (L2), the physical layer, which is discussed separately in the next section. On the transmitter side, one common layer needs to oversee the allocation of incoming service data units (SDUs) to each RPU (as discussed above). To maximize the number of tasks that can be executed in parallel, this needs to be located as high up in the radio protocols as possible. An ideal candidate would be the higher part of the Packet data convergence protocol (PDCP) layer, after sequence number (SN) allocation but before other functions such as security and header compression. This would allow these other functions to be performed in parallel on each RPU while allowing the receiver to re-order the SDUs coming out of the RPUs. An example is depicted in figure 7. Note that, as explained above, whether a part of a sublayer of L2 being hosted on one RPU becomes visible to the receiver is likely to depend on whether the RPUs operate on common or segregated resources.

Figure 7. Two stacks for the user plane in 6G



To facilitate the provision of very high bitrate services in RPUs in the FPS, all the processing should be as fast as possible. While 5G radio protocols considered implementation complexity and pushed real-time functions as close as possible to the physical layer to maximize offline processing [1], there is an opportunity for the FPS to be even more friendly to high bitrate processing by dropping some optimizations. We will now explain how this could be achieved.

L2 processing mainly consists of processing headers: generating headers on the transmitter side; interpreting and taking actions according to the headers on the receiver side. Therefore, the headers should be as simple as possible, specifically they should be in fixed positions and of fixed length even if that means increasing the overhead by a few bits as it hardly matters for very high bitrate services. Fixed headers would significantly speed up processing of the headers and would even allow hardware acceleration.

Examples for L2 header design for fast processing:

- Fixed SN length for all layers where needed, e.g., fixed 32-bit SN (full COUNT) for PDCP
- Fixed RLC unacknowledged mode (UM) header with SN and segment offset (SO) always present
- Fixed logical channel identifier (LCID) length and fixed size length field for MAC.

Furthermore, to speed up user plane processing in the FPS, dynamic MAC control elements related to control functions should be limited to the APS to guarantee quick parsing of PDUs in the FPS. Other simplifications in the FPS could also include not having any ARQ function at RLC.

Physical layer

In terms of RPU design, a key question is whether the physical layer processing is part of the RPU or not.

Physical layer processing is part of the RPU. In this approach, each individual RPU generates transport blocks (TB) matching the allocations from the physical layer and maps the output bits to its designated radio blocks (RB) without interaction with the other output bits and TBs from the other RPUs. As a result, RPUs, including HARQ (hybrid ARQ) and their physical layer, can run in parallel. This option resembles carrier aggregation architecture in 4G and 5G. Using different transmit parameters such as modulation coding schemes (MCS) and RBs per RPU is possible, but they would increase signalling complexity and overhead. This approach would tie an RPU to a physical resource and would not allow an RPU activation scheme fully based on load (as described earlier) and thus, is not the preferred option.

Physical layer processing is NOT part of the RPU. In this approach, the physical layer processing works across RPUs, at least within one cell or component carrier. Based on the allocations determined at the physical layer, each RPU that is available (with non-empty buffer) is requested to deliver one MAC PDU. Each RPU can map the output bits to a deterministic location of the MAC-PHY interface memory without interacting with other RPUs. The TB is then passed to the physical layer for processing. In this case, it seems more natural to assume common HARQ processes and buffers for RPUs within the same cell or component carrier. By decoupling RPU from the physical layer, this option would give more freedom to the management of RPUs at L2, and allow parallel processing at L1. Thus, this is the preferred option.

The analysis above is summarized in Table 1 below.

Table 1. Summary on placement of physical layer processing relative to RPU

Physical layer processing	Flexible retransmissions	Load-based RPU activation	Decoupling of parallel processing between L1 and L2
Part of RPU : one-to-one mapping between TB and RPU	**No** L1 retransmissions and RPUs are tied together	**No** RPU activation needs to be linked to TB (re)transmissions increasing complexity and constraining power saving opportunities	**No** Parallel processes at L1 and L2 must be aligned
Outside of RPU : RPUs and TBs need not be coupled	**Yes** L1 retransmissions can be decoupled from RPUs	**Yes** RPU activation can maximise power saving opportunities	**Yes** As long as the total bit rate is the same, the number of parallel processes can be different between L1 and L2

Conclusion

While the use cases, deployments scenarios, and verticals that 6G needs to address are essential topics to discuss to shape the future of 6G, the design principles of the radio protocols addressing them must be tackled early on to secure that future. This paper has highlighted the limits of the one stack approach that has prevailed in earlier generations. Instead, we are proposing a paradigm shift in how radio protocols are designed by relying on a dual stack approach, with a first stack hosting control plane functions and optimizations for low bitrate services, and a second stack optimized for parallel processing with the notion of radio processing units or RPUs. We also suggest decoupling RPUs from the physical layer to minimzie dependencies between L2 and L1, and facilitate parallel processing in both layers. With those principles in place, 6G radio protocols could be less complex and less power consuming than earlier generations but without sacrificing flexibility or performance.

Abbreviations

AI	Artificial intelligence
APS	Anchor protocol stack
ARM	Advanced RISC machine (originally Acorn RISC machine)
ARQ	Automatic repeat request
CISC	Complex instruction set computer
CN	Core network
COUNT	PDCP data PDU counter
CP	Control plane
CS	Circuit switching
FPS	Fast protocol stack
GPRS	General packet radio service
GSM	Global system for mobile communication
HARQ	Hybrid ARQ
L1	Layer one
L2	Layer two
LCID	Logical channel identifier
LTE	Long-term evolution
MAC	Medium access control protocol
MCS	Modulation and coding scheme
ML	Machine learning
MTC	Machine-type communications
NB-IoT	Narrowband internet of things
NR	New radio
PDCP	Packet data convergence protocol
PDU	Protocol data unit
PHY	Physical layer protocol
PS	Packet switching
QoS	Quality of service
RAN	Radio access network
RB	Radio blocks
RedCap	Reduced capability
RISC	Reduced instruction set computer
RLC	Radio link control protocol
RRC	Radio resource control
RPU	Radio processing unit
SDAP	Service data adaptation protocol
SDU	Service data units
SN	Sequence number
SO	Segment offset
TB	Transport blocks
UM	Unacknowledged mode

References

[1] Holma H., Toskala, A. and Nakamura, T., 5G Technology: 3GPP New Radio, Wiley, 2019.

[2] Next G Alliance, Next G Alliance Report: Roadmap to 6G, Alliance for Telecommunications Industry Solutions, Washington DC, 2022. https://www.nextgalliance.org/white_papers/roadmap-to-6g/

[3] Mediatek, 6G Vision Whitepaper, Mediatek, 2022. https://d86o2zu8ugzlg.cloudfront.net/mediatek-craft/documents/MediaTek-6G-Vision-White-Paper-EN0122.pdf

[4] NTT Docomo, 5G Evolution and 6G: white paper, NTT Docomo Inc. 2023. https://www.docomo.ne.jp/english/binary/pdf/corporate/technology/whitepaper_6g/DOCOMO_6G_White_PaperEN_v5.0.pdf

[5] Nokia, Nokia's vision for the 6G era, Nokia web site, viewed Jul 19, 2023. https://www.nokia.com/about-us/newsroom/articles/nokias-vision-for-the-6g-era/

[6] Next G Alliance, Next G Alliance Report: 6G Technologies, Next G Alliance, Jun 2022. https://www.nextgalliance.org/wp-content/uploads/dlm_uploads/2022/07/TWG-report-6G-technologies.pdf

[7] Aoudia, F. A., Hoydis, J., Valcarce, A. and Viswanathan, H., "Toward a 6G AI-Native Air Interface," Nokia Bell Labs white paper, Mar 2021. https://www.bell-labs.com/institute/white-papers/toward-6g-ai-native-air-interface/#gref

[8] Kanugovi, S., Montag, M., Viswanathan, H. and Ziegler, V., "Security and trust in the 6G era," Nokia Bell Labs white paper, Aug 2021. https://www.bell-labs.com/institute/white-papers/security-and-trust-6g-era/#gref

Nokia OYJ
Karakaari 7
02610 Espoo
Finland
Tel. +358 (0) 10 44 88 000

631701 (August) CID213456