NOKIA Bell Labs

Technology innovations for 6G system architecture

White paper

G. Kunzmann, B. Masood Khorsandi, V. Ziegler, V. Raisanen,
E. Ozturk, T. Kolding and C. Mannweiler

Contents

Introduction

Approximately every ten years, the opportunity presents itself to define the next generation of mobile networks. While the roll-out of 5G systems has successfully evolved, with currently more than 200 networks being operational worldwide, the journey of 5G evolution will continue for another eight years with innovations coming as evolutionary steps under the umbrella of "5G Advanced" [1]. This will comprise solutions to enhance coverage, user experience, expand beyond connectivity and operational excellence.

At the same time, 6G research is in full swing together with strong momentum from pre-competitive joint research initiatives such as the Hexa-X European 6G flagship project [11] within the scope of the Horizon 2020 program, Next G Alliance in the US [9][10], China's IMT-2030 (6G) Promotion Group [4], Beyond 5G Consortium in Japan [5], and the Korean government's (MSIT) 6G R&D implementation plan [6].[1] 6G demonstrates its difference from previous generations by recognizing from the beginning the need for defining key value indicators (KVIs) such as sustainability, trustworthiness and digital inclusion to drive key challenges of research, as well as to enable various use case families of the 2030s. In other words, considerations of both "societal pull" and "technology push" drive the 6G research journey in scope and substance. Novel use case families for communications in the 2030s will include, among others, immersive telepresence, massive twinning, collaborating robots as well as trusted and specialized 6G subnetworks [14][8].

Six key technologies will enable the essence of 6G including new spectrum bands and technologies, artificial intelligence (AI) and machine learning (ML) techniques, network-as-a-sensor, extreme networking with ultra-low latency, cognitive automated and specialized architecture, as well as new concepts to assure security and privacy [2]. Finally, one of the most distinct key enablers will be a new 6G architecture as described within the scope of this white paper to help consolidate and integrate 6G technology enablers. 6G offers the opportunity to introduce major innovations without full backward compatibility.

The 6G system and the associated architectural decomposition into functional focus areas are embedded into broader architectural themes of transformation. The 6G system includes the UE, access and core network functions. It includes novel functions for 6G radio access, allowance for new spectrum and accommodations for new use cases. It will also include functions from previous generations (4G and 5G) as well as non-3GPP access technologies. The distributed heterogeneous cloud ("Het-Cloud") provides the open, disaggregated, and distributed execution environment for the network functions [3].

The essential network fabric is the secure, seamless, cross-domain network that assures the needed connectivity for humans, machines, industrial applications and disaggregated computation. From a holistic point of view, 6G architecture forms a "network of networks", which is heterogenous from a technology and business model perspective. We define the network of networks in terms of critical 6G sub-networks, ubiquitous RANaaS, as well as options for "360° connectivity" by means of novel multi-connectivity mechanisms on different layers. Its services can be consumed seamlessly, and it can be operated in an efficient and highly automated manner.

6G architectural principles will be defined to support AI/ML, service-oriented orchestration with closed-loop automation and intent-based management, and the adoption of cloud-native principles like service-based design, disaggregation and openness, platform agnosticism, and continuous integration and delivery.

Within the scope of this white paper, in Section 2 we will explore aspects of the 6G vision, use cases and enablers for 6G architecture as currently envisioned. Functional goals and decomposition of the 6G architecture will be described in section 3 with special emphasis on RAN-Core architectural simplification and evolution.

1 An overview of global 6G initiatives can be found in Vetter [7].

Use cases and innovations

In this section, we will review the current developments related to and impacting the 6G architecture, as well as initiatives focused on shaping the future 6G system. Nokia expects 6G to launch commercially around 2030 [2], and, soon after, 6G is expected to dominate mobile devices, network sales and mobile service usage. To design a 6G system that realizes this vision and plays a key enabling role for society in the 2030s, a first step was to define relevant use cases. The defined 6G use cases are not only a simple extension of 5G use cases, but also outline new uses aimed at a more inclusive (in the geographical as well as the societal sense) and more sustainable use of the technology. The comprehensive list of 6G use cases can be found in Hexa-X deliverables D1.2 and D1.3 [13] [14]. As an example, we discuss two examples, the sustainable development use-case family and the immersive telepresence for enhanced interactions use-case family.

Sustainable development use-case family

To address the growing concern over environmental sustainability, the 6G system will need to provide a meaningful impact on the environment, for example, by providing solutions helping to meet the United Nations Sustainable Development Goals (UN SDGs) [16] or helping verticals to reduce their ecological footprint. One of the use cases from this family is “network trade-offs for minimized environmental impact”. This use case family is based on the idea of providing standardized indicators for energy consumption and tuning of the network. The latter capability would encompass the selection of appropriate paths and network management considering criteria such as paths powered by renewable electricity or avoiding paths involving non-circular material usage. The indicators would enable measurable trade-offs between Quality of Experience (QoE) and environmental footprint to better meet the targets for minimizing environmental impact.

Immersive telepresence for enhanced interactions use-case family

This use-case family is about users being present and interacting anytime anywhere, using all their senses, if so desired. It enables humans to interact with each other as well as physical and digital things around them. This calls for new sensing solutions to compose multi-layered maps of our environments. Radio sensing using the mobile communication network as a sensor has the potential to become an essential component of the solution [15]. The use cases of this family require a system architecture in which extreme low latency can be provided. For example, “fully merged cyber-physical worlds” uses mixed reality and holographic telepresence to appear as though one is in a certain location while actually being in a different location. This makes it possible for users to communicate with distant persons with a quality of interaction very close to physical presence. Another example is an “immersive sport events”. With the projected capabilities of a 6G system, it will be possible to motion-capture actual games in real time to create a digital twin of the whole game, which can be experienced live from any angle, by hundreds of millions of people worldwide.

6G systems will be characterized by a collection of technologies, use cases and services running on top of a heterogeneous hardware infrastructure and an enabler platform, requiring rethinking of how the system is architected, designed and deployed. These innovations include, for instance, extended reality (XR) artificial intelligence (AI) and machine learning (ML), to name a few of these innovations.

Extended reality (XR)

Building on VR and AR, XR will develop in many different industries such as entertainment, medicine, science, education and manufacturing. Such techniques require not only huge data throughputs, low latency and positioning, but also, specific device features like compute and battery, which can be implemented in specific devices such as lightweight glasses. In addition to being multi-sensory, these features make a holographic representation of objects and people or digital twin (DT) possible. For example, you could attend, remotely, a meeting with your colleagues all being present in the same virtual meeting room hearing discussions, seeing each other, and even smelling as if you were physically in the room. To support this technology, the 6G system should be able to provide a digital representation of the physical world, based on sensors and AI technologies.

Artificial intelligence (AI) and machine learning (ML)

AI/ML technologies are the essential ingredient for 6G architecture. These technologies will help unlock the full potential of 6G systems, empowering them with new network automation capabilities, boosted performance and enhanced energy efficiency. Additionally, new services and applications employing AI/ML will impose new requirements for the 6G architecture.

New X-as-a-service innovations

There are various other emerging use cases that may require resource-constrained devices and are based on data-driven decisions, which will require new services such as compute-as-a-service (CaaS), general-purpose AI-as-a-service (AIaaS), and AI-assisted vehicle-to-everything (V2X). Each of these services is, in their turn, enabled by key technologies that are building blocks of the future 6G architecture.

The vision defined for 6G is built on the interactions between three intertwined worlds, physical, digital and human, and these interactions are driven by three core values, sustainability, inclusiveness and trustworthiness.

The performance goals of the 6G system need to go beyond what 5G can do to address new use cases. In designing the 6G architecture, it is necessary to have key performance indicators (KPIs) that can describe the increasing peak data rates (including data rates achievable at the cell edge), the density of connections, traffic capacity and location accuracy with the required granularity and level of accuracy. Furthermore, in specific use cases that require extreme performance, the focus will shift towards new end-to-end KPIs that can measure performance goals such as dependability and determinism, service availability, affordable coverage, and network energy efficiency [13] [14].

Moving towards 6G, it is necessary to expand the fundamental network design paradigm from performance-oriented design to both performance- and value-oriented design to fully embrace the 6G vision. This requires a new class of evaluation criteria such as key value indicators (KVIs) that measure sustainability and trustworthiness, which must be understood, developed and adopted in the system architecture design evolution to 6G.

For 6G to become the key enabler for making other sectors more sustainable, it needs to be a sustainable technology in itself. There are several factors that must be considered when designing the 6G architecture for a sustainable future mobile network. As an example, 6G should be designed to reduce absolute network energy consumption and associated carbon emissions, as well as to reduce the energy consumption per transported bit. Based on current technologies, the continuous traffic growth delivered by telecommunication networks will lead to an increased need for energy. We can counter this by adopting advanced technologies that are intrinsically more energy efficient and by optimizing the energy performance of the overall network. For example, embedded energy monitoring systems and sites management could optimize and adapt energy consumption to traffic flows and system activities. This would allow us to optimize the end-to-end energy efficiency behaviour. New techniques of energy efficiency should be included in all the network components to have zero consumption at zero load.

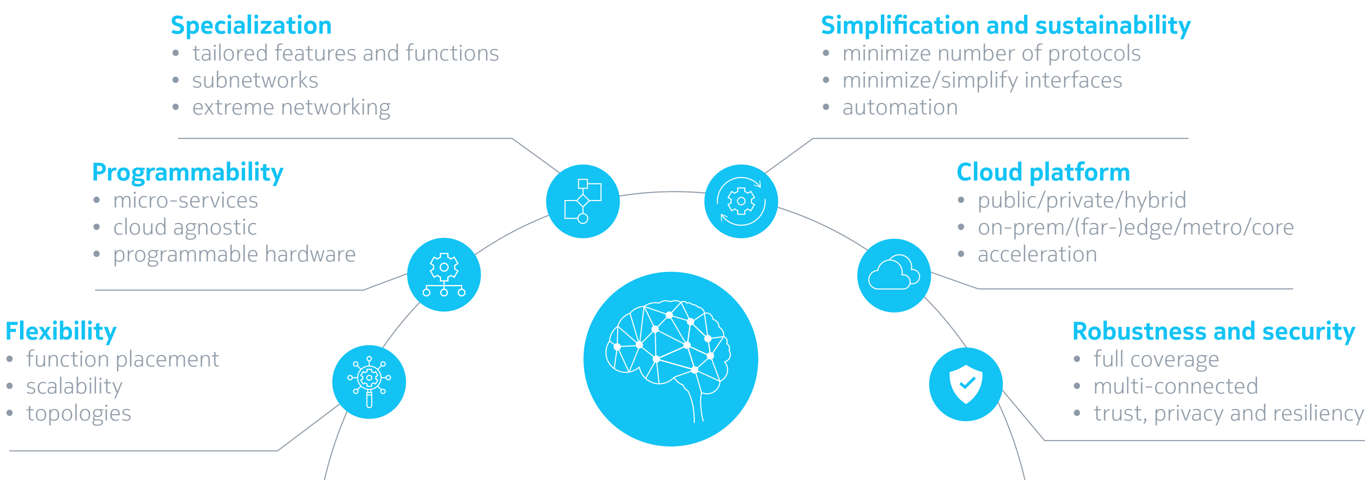
Trustworthiness is one of the main 6G research challenges. Security considerations must encompass all aspects of cybersecurity: resiliency against attacks, preservation of privacy, and ethical, safe application of automation to network operations and applications. A realistic approach to this trustworthiness challenge is to acknowledge that all security measures come with costs, such as usability, agility or swiftness. Thus, a balance is required. Providing the necessary elements for the evaluation of this balance is the key goal for a secure 6G architecture design.

Towards a 6G architecture

6G architecture design goals

As introduced in the previous section, the ecosystem is expected to have new requirements, both KPIs and KVIs, for the 6G architecture to support a set of new use cases, while still supporting existing use cases in an optimal manner. To address these requirements, the 6G architecture is expected to evolve in six key areas:

Figure 1. 6G architecture design goals



Flexibility: The 6G architecture is expected to be more flexible in various dimensions. First, the architecture must work for large-scale wide area network deployments as well as for extremely local on-premises and personal area networks. Second, in terms of function placement, the 6G architecture is expected to achieve a wide variety of latency targets and meet other requirements in a dynamic manner. For example, to support XR services, it will need to achieve ultra-low latency for specific users in constrained geographical areas and time windows. Ability to dynamically scale, for instance, to meet changing network loads, is mandatory.

Specialization: Given the variety of use cases and deployments, the 6G architecture also needs to be able to provide tailored features and functions. For example, a specific 6G sub-network, a lightweight sensor network, or some extreme networking use cases may require a specific set of functions, while other features can be omitted as they are not needed in this type of network.

Robustness and security: Users in vertical industries expect that the network service is provided in a robust, resilient and truly ubiquitous manner allowing for multi-connectivity where applicable and required. Also, we must leverage the best from different worlds, for example, facilitating the co-existence of a macro network architecture with full mobility and roaming support while seamlessly interworking with an architecture for a short-range, cell-less THz network and a non-terrestrial network architecture (NTN). A further core goal for security is to meet the strong expectations around trust and privacy.

Cloud platform: The ongoing shift to hosting of network functions in cloud platforms will continue and become broader. Deployments will, to a large extent, move from specialized telco cloud platforms to generic public, private, or hybrid clouds that can be located on-premises, in the (far) edge or in a central site. 6G will provide a uniform orchestration interface for service management of distributed clouds, complemented with segment specific abstractions, e.g., for the 6G RAN. This allows for use of specialized capabilities in the cloud nodes, such as, hardware acceleration.

Programmability: Implementations will also achieve a new level of programmability, like hardware-independent programming languages such as P4, and will need to run on any cloud platform. Serverless services and functions from multiple vendors will become even more easy to integrate through open and service-based interfaces and a more modular system design. This will allow deployment of the network or instantiation of network slices with those, and only those, functions and services that are required.

Simplification and sustainability: As we expect more and more from our networks, at the same time we also must investigate opportunities for simplification of the architecture. The complexity and number of functions in the system has grown over the last generations. While more powerful zero-touch automation capabilities offer one means to cope with system complexity, the introduction of 6G also offers the opportunity to re-visit the architecture and clear out, re-design or merge functions, introduce a service-based approach also for the control plane between the Radio Access Network (RAN) and the core network jointly with a distributed Non-Access-Stratum (NAS) interface, simplify signalling procedures, etc. As another example, simplification can be achieved by a tailored protocol stack, allowing for an efficient integration of customized 6G subnetworks. Simplification and providing more flexibility and dynamicity in function placement will also help to achieve sustainability goals by reducing the amount of signalling and energy consumption. A clear commitment to “design-to-sustainability” will be important for 6G.

Artificial intelligence and machine learning technology is expected to be an integral part of the 6G architecture. It is a technology that is needed to achieve the vision of a truly cognitive network that adapts itself to a variety of scenarios and deployments. AI/ML is expected to play a key role in powering the automation and optimization of the network, as well as in increasing the security of the system. The extensive use of AI/ML requires the architecture to incorporate functions and interfaces for large-scale data collection, processing and distribution, training for model refinement, and updating inference models within functions.

Besides the listed architecture design goals, there are additional aspects to be considered on our path to a 6G architecture, such as migration from 5G, deployment options to be supported, and identifying an appropriate standardization approach for 6G. The architecture needs to be future-proof, agile, cater to new and unknown applications, and consider various technology trends expected in the 6G era. It is also essential that the 6G system be able to interwork with, at least 5G, and ideally, previous generations.

Developing a new generation is an opportunity to explore the system architecture with a clean slate, i.e., start over without prejudice and with full freedom in the architecture design. Thus, some of the 6G architecture goals are expected to require innovations introducing non-backward compatible changes. At the same time, 6G will likely also adopt tested and proven concepts and functionalities from 5G. Some of the 6G drivers requiring non-backward compatible changes, compared to 5G, involve:

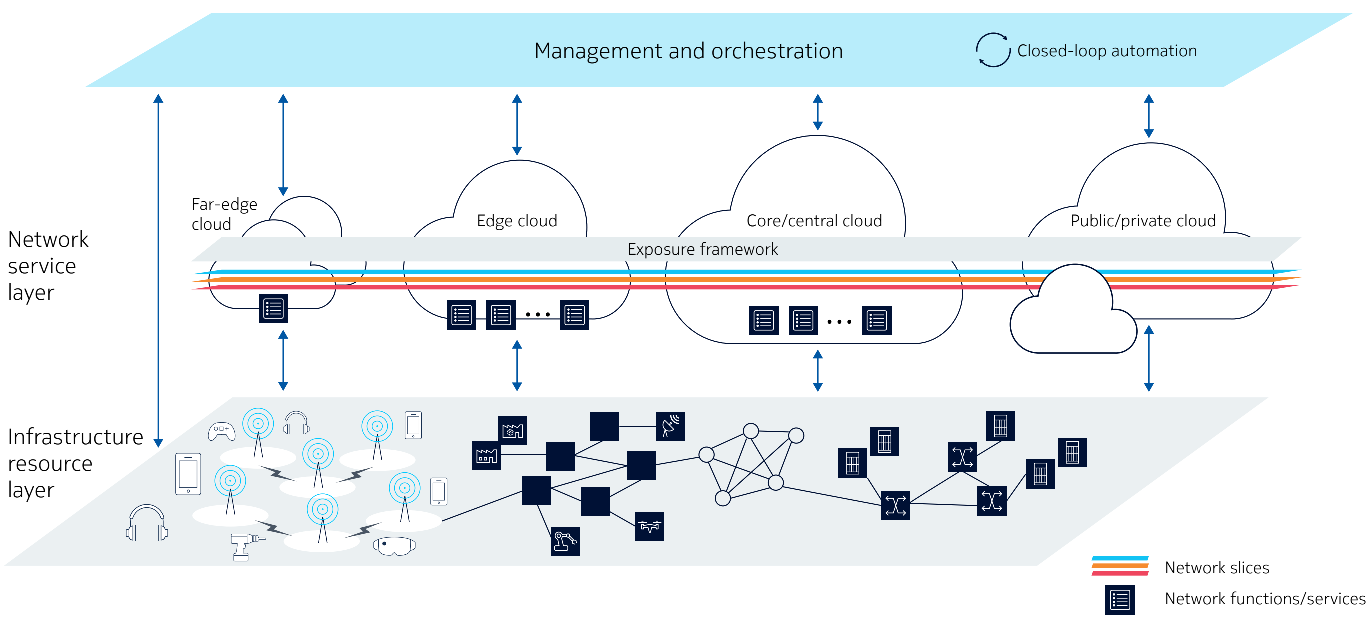
- Designing a new non-access stratum (NAS) protocol stack and RAN-core interface
- Designing a new user plane (UP) architecture to address the needs of new use cases and security requirements
- Supporting passive Internet of Things (IoT) devices requiring light-weight system procedures

- Enabling a massive increase of data collection for data analytics, AI/ML model training, and AI/ML model transfer
- Supporting new application trends, such as the Metaverse [20].

End-to-end 6G architecture view

Figure 2 shows the envisioned end-to-end (E2E) 6G architecture. The infrastructure resource layer provides the physical resources to host the network functions and services. The layer includes all the switches, routers, transport links, data centers, cloud infrastructure and radio equipment, including non-virtualized radio functions like radio units (RU), distributed units (DU) and the base stations. The network service layer is expected to be fully cloud-based, comprising different clouds that can be grouped into far edge, edge, core/central, and public/private clouds and can be provided by different stakeholders. The far edge will be characterized by a heterogeneity of devices with a wide variety of software and hardware technologies such as personal and IoT devices.

Figure 2: End-to-end 6G architecture view



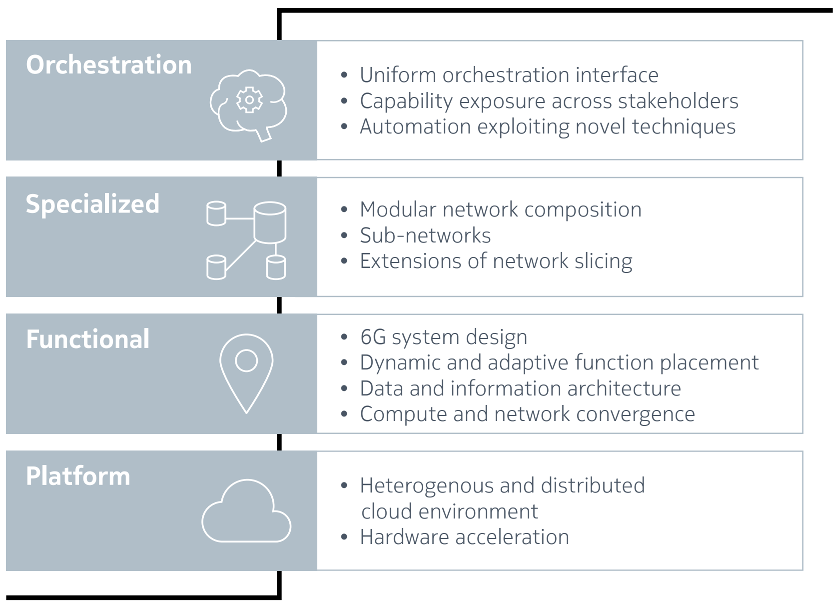
An exposure framework is envisioned to establish a control channel between the multiple clouds, enabling seamless interoperation and networking. By having all network functions, operations and applications implemented as services, the 6G architecture can be softwarized, intelligent and efficient. The applications running over the 6G systems will request communication and computation services with specific QoS and QoE needs that can be delivered by a dedicated network slice.

The management and orchestration domain, depicted on the top of Figure 2, spans all layers of the 6G system. It is gradually moving toward increasing the levels of automation and utilizing fully automated closed-loop control. This is supported by the parallel adoption of advancements in AI/ML technologies.

Overview on 6G architectural innovations

In this section, we discuss some of the architectural innovations that should be considered to meet the above mentioned 6G design goals. These innovations span all layers of the 6G system and are comprised of features that include the 6G hardware platform, general mobile network functions, specializations for specific use cases and requirements, and orchestration layer components, as depicted in Figure 3. After a first high-level description of the architectural innovations, we will present some innovations in more detail.

Figure 3. 6G architectural innovation



From a **platform** view, we envision a heterogeneous and distributed cloud environment, i.e., the 6G network may spread across multiple clouds at different locations and may involve multiple stakeholders. Cloud capabilities, such as hardware acceleration, will be essential components that will be integrated, e.g., to support deep learning processing and cloud-native implementation of lower layer RAN functions.

At the **functional level**, we can design a new functional split within RAN, between RAN and core network functions, with the aim to allow for an optimized UE-RAN-Core architecture design. This will simplify related procedures and consolidate similar functionalities. Possible solutions range from simply enhancing the existing interfaces to completely removing the separation of domains — or something in between. Dynamic and adaptive placement of functions may best address the trade-offs between network performance, complexity, energy consumption and, most importantly, service and consumer demands.

At this same level, we also expect changes to the massive amounts of data and information that are collected, stored and exposed, for example, to AI/ML-powered analytics functions at all layers of the mobile network. Data are created in and collected from multiple areas and layers of the network and can be delivered to where they are needed. Therefore, we see the need for a common exposure, registration, discovery and delivery framework for a wide range of different data types, such as near-real-time streamed data, state data for services in various registrars, and accumulated data for training ML models. Also, a convergence of networking and computation is under investigation. In this context, ensuring privacy and providing trusted solutions is essential.

As discussed earlier, the 6G network is going to support a vast range of diverse and **specialized** purposes and deployments, such as, service availability on local vs. global scale, enterprise vs. CSP deployments, best-effort vs. mission critical traffic, and so on. This diversity comes with specific requirements and will utilize specialized capabilities of the network. One of those capabilities is network slicing, which was introduced in 5G but must be extended with new concepts such as dedicated SW stacks and dedicated HW, to meet 6G requirements like the complete separation of sensitive traffic flows. Additionally, extreme performance requirements will need to be achieved in specific networking contexts such as body area networks or robot networks that could be part of the 6G network. We also refer to these types of networks as 6G subnetworks or ‘in-X subnetworks’ [17]. Our vision is to develop a modular architecture, which can easily be composed depending on the target deployment, its use case(s) and their specific needs.

Finally, 6G will provide uniform **orchestration** for the domains in a 6G system, as well as concurrent abstractions referred to earlier. Multiple network domains need to be orchestrated to compose and support an end-to-end communication service that may span multiple stakeholders and administrative domains such as CSP, enterprise and home domains. Moreover, service orchestration needs to support various business constellations relevant to 6G, including capability exposure across stakeholders corresponding to monetization models and end-to-end orchestration of far edge resources and 6G subnetworks. Automation is required to achieve 6G performance targets for orchestration. Thus, architecturally consistent use of techniques like intent-based management, closed-loop automation and AI/ML will be important technologies.

In the subsequent section, we will cover selected innovations and implications in more detail.

Platform for system abstraction

Increasing virtualization and cloudification in communications networks can benefit from a common computation, storage and delivery platform, creating an abstraction layer for applications and services to enable their fast development, easy deployment and efficient operation.

To support novel deployment options as well as use cases and their requirements, the platform can utilize different cloud technologies that provide different sets of features and capabilities, hardware types, high availability characteristics, and supported computing paradigms such as containers or serverless. At the same time, certain services may not be fully cloud-agnostic and require specific capabilities such as hardware acceleration. Thus, they cannot be placed in any arbitrary cloud. Consequently, defining the required cloud capabilities of the deployed services and a cloud capability discovery are needed to optimally select cloud instances where the services can be placed. Moreover, for an optimal placement of the functions, the cloud selection also needs to consider the target service, such as its QoS requirements and where the service will be consumed. The placement decision also needs to have overall network optimization goals in mind.

Cloud deployments can leverage cloud-native technologies to further improve the communication; for instance, the adoption of a service-based architecture (SBA), where functionalities are delivered via independently maintainable inter-connected network functions (NFs) that provide a set of well-defined services communicating through service-based interfaces (SBI) to enable flexible and scalable deployment and fast-to-market improvements. Currently, SBA principles have already been employed in the 5GC domain for control plane functionality and in the network management domain.

6G system design

Functional

- 6G system design
- Dynamic and adaptive function placement
- Data and information architecture
- Compute and network convergence

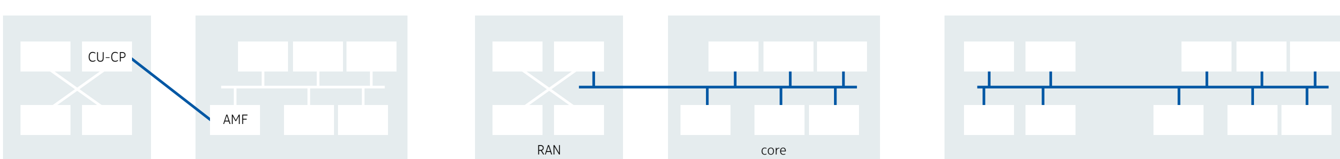
On the functional level, one of the major research areas under consideration are new functional splits within the RAN and between the RAN and the core network. From the RAN point of view, it can be advantageous to deploy some of the functionalities in a more centralized manner. For example, it may be more efficient to centrally perform certain tasks for a group of cells instead of doing the computation at each base station (more simple base station design) or being able to optimize network parameters for larger geographical areas (while base stations will be deployed more densely).

This trend is already visible in 5G. For instance, the central unit (CU) hosting the latency-tolerant higher layers of the radio protocol stack can cover a large service area and can be deployed in the edge or regional cloud. It works in combination with distributed units (DUs) that deal with the latency-sensitive lower layer aspects and are placed in close vicinity to the cell sites and the radio frequency (RF) unit. In this way, one CU can serve a large set of DUs.

In 6G, the modularity of the RAN and leveraging of centralized processing (e.g., with centralized function deployment being dynamically adjusted based on service requirements and network status) is not only expected to continue, but will be extended to meet the needs related to massive data collection and processing that will be an integral part of next-generation communications. Centralized functions will help realize operation, monitoring and control within the RAN as well as coordinating among RAN nodes in an ever-more optimized manner.

In addition to moving RAN functions closer to the network centre, certain RAN and core network functions will need to be pushed towards the network edges and deployed in a distributed manner. The hierarchical processing that exists in current 5G networks will not support the demanding new use cases that are defined (and yet to be defined) for 6G. Upon analysis of these use cases and their requirements, such as service latency and real-time constraints, certain RAN and core features may need to be deployed at full or almost full capacity in a localized manner, closer to the UEs. This means that the current RAN logical split and RAN and core functionalities need to be redefined so that they can be dynamically deployed in the same physical sites and share supply models and infrastructure mechanisms to support the defined use cases.

Figure 4. RAN – core redesign



a) Current N2 interface b) Service-based RAN-core interface c) Service-based RAN and core interface

As of today, the RAN and the core network control plane are connected via the N2 interface between the CU-control plane (CU-CP) and access and mobility function (AMF) as depicted in Figure 4(a). Due to the RAN and core network functions being co-located, in this example, functions beyond the CU-CP and AMF are likely to end up being deployed in the same cloud site, while signalling remains bound to the RAN-core interface.

To address this trend of RAN and core components being co-located, several potential solutions are currently being investigated. For future 6G systems, service-based architecture principles are expected to be applied to the RAN-core interface by defining a service-based interface between RAN and core as in

Figure 4(b). This will enable a direct interface between the RAN and multiple NFs within the new 6G core.

Adapting the RAN-core interface to be service-based is not sufficient to exploit the full potential of service-based architecture principles if the non-access stratum (NAS) protocol terminates in a single point in the network. This would mean that all the signalling messages sent and received by the UE would still be using point-to-point communications, thus not fully exploiting the service-based interface between RAN and core. To leverage the full potential of SBI, distributed NAS termination needs to be considered. It would enable signalling between UE and appropriate network functions without having to cross a single point of termination every time. This design is expected to come with a range of benefits like easy and flexible adding of new functionalities to support new services, reduced signalling loads at the single point of termination, reduced signalling latency due to direct communication, and strengthened security for each termination by keeping them independent (e.g., avoiding a single point of failure) so as to enable NAS security between UE and each NF.

For a further step towards a fully service-based RAN-core signalling, a complete opening among RAN and core functions could be considered, as shown in Figure 4(c), where all RAN and core functions communicate directly with each other. Functionally, RAN and core services would still need to perform standardized and predictable operations needed for their separate domains, but core and RAN functions can provide a larger range of services that can be exposed to the other domain directly via SBIs. This needs to be evaluated on a case-by-case basis as today's clear separation between the RAN and the core functions has a valid justification in its separation of concerns. Applying the SBA principles inside the RAN comes with its own challenges as many of the physical network designs found in the RAN are not well suited for a service-based design. Some issues, like latency, would need to be considered.

Data and information architecture

Functional

- 6G system design
- Dynamic and adaptive function placement
- Data and information architecture
- Compute and network convergence

Big data analytics have started to enter the mobile communication systems. For example, 3GPP SA2 specifies, from Rel-15 onwards, a network data analytics function (NWDAF) that can provide statistics and predictive analytics for a growing set of use cases [18]. In the management and orchestration domain, SA5 specifies a management data analytics service (MDAS) [19], and analytics functionality is also proposed for the RAN. To derive the requested analytics reports, the analytics services may collect data from a variety of data sources such as events from other NFs, performance data from the OAM domain, so-called minimization of drive test (MDT) measurements related to a specific UE via OAM, and external parameters from application functions (AF) via the network exposure function (NEF).

The NWDAF also comes with the capability to (re-)train the AI/ML model(s) it is employing, where again, it needs to collect a large among of training data from various data sources. As a consequence of the increasing use of analytics functions in other NFs and the need for highly accurate analytics, it is essential to extend data collection to cover more parameters and events, utilize more data sources, and store such data for longer time windows. A dedicated data and information architecture is envisioned for 6G to collect, store and expose this data across the domains. Different to the subscription-notify approach currently specified for the core network, a publication-subscription pattern or event stream approach might be better suited to handle the immense data to be exchanged in the data and information layer.

Compute and network convergence

Functional

- 6G system design
- Dynamic and adaptive function placement
- Data and information architecture
- Compute and network convergence

In addition to the above innovation, requested network services will no longer be limited to pure connectivity. Rather, many use cases will also require compute and storage capabilities, for example, extended reality (XR) and digital twinning. It is essential to optimize user-plane applications placement based on the specific networking, compute and storage demands of the application and its users. To determine the optimal application placement from an overall system perspective, it is not sufficient to find the closest edge cloud or data center but also to consider other needs of the application, including available computing capacities, possibly with specific hardware requirements, such as available GPU- or AI-acceleration. Convergence of networking and computing capabilities is likely to happen in 6G and details of such an architecture are under study. One sensible approach is to make the system aware of additional information related to the computing and storage resources available within the network and towards the AFs. User plane functions may then be placed on the same rack in the same data center and close to the user to reduce latencies, but also to avoid unnecessary intra-data-center traffic, thus, contributing to the sustainability efforts desired for the 6G ecosystem.

Domain automation

Orchestration

- Uniform orchestration interface
- Capability exposure across stakeholders
- Automation exploiting novel techniques

Advanced domain automation functionalities are needed to support a uniform northbound orchestration interface for resource domains. Domain automation techniques are employed for cloud RAN evolution during the 5G era for core and regional edge domains. Automation for the emerging far edge cloud platforms addresses the low aggregation level and high availability challenge encountered in extending cloud-native principles to the network edge. Particularly at the edge of the network, it is important to employ automation to the management of both the telecommunication functionalities and the cloud platforms. End-to-end 6G services may also include resources in the far edge domain beyond the traditional mobile network. Closed-loop automation is expected to have an important role in domain automation in complementing intent-based management and facilitating a uniform orchestration interface for domain resources as described in the Hexa-X project. Machine learning will be important for closed loop automation, consequently, it needs a data architecture supporting the various automation use cases as well as AI/ML pipelines and model management frameworks for online and offline learning. In view of the fundamental role of AI/ML in 6G, techniques such as federated learning are expected to be important to support 6G performance levels while also catering for the variability of edge environments in a resource-constrained inference environment.

Conclusion and outlook

New KPIs and, in particular, KVIs such as sustainability and trustworthiness play an essential role in the development of the 6G system. Further, 6G must create added value for all stakeholders, including mobile network operators, enterprises, vertical industries and end users. All the players expect support for different use cases and requirements and can benefit from an architecture configuration and deployment that can be tailored for their specific needs. At the same time, new enabler technologies need to be embraced to meet those requirements and related design goals. Nokia has identified six key design areas for 6G [2], the 6G architecture being the most distinct area. In this whitepaper, we outlined selected components and considerations for the 6G architecture.

Innovations described in this white paper range upwards from the platform level through the functional level to a uniform orchestration interface, and all concepts studied are contributing to the defined goals of the 6G architecture. In addition to those specific concepts, it is essential to keep the end-to-end 6G system in mind and evaluate the individual solutions from a holistic 6G point of view. As an example, several of the concepts being discussed will require a new clean-slate 6G architecture to exploit their full potential. As such, migration from, and interworking with, previous generation systems must be considered as well as consideration for an appropriate standardization approach for 6G.

Across the globe, several 6G initiatives have already started research toward the next mobile network generation. In Europe, Nokia is leading the flagship project Hexa-X [11]. The main goal is to provide the 6G vision and intelligent fabric of technology enablers to connect the human, physical and digital worlds. Nokia plays the main role both in coordinating the key industry stakeholders and technology providers and contributing to almost all technical topics ranging from RAN technologies, localization and sensing to AI-driven communication and computation, as well as enablers for 6G architecture and management and orchestration. Nokia is contributing to architectural enablers [12] such as network programmability and dynamic function placement, as well as contributing to the proposal of the first draft of end-to-end 6G system view architecture [14]. Nokia plays a leading role in competitive initiatives in North America such as the Next G Alliance initiative [6, 7] that were launched in 2020 to advance North American mobile technology leadership. Nokia is a founding member and serves in leadership roles within the Next G Alliance Steering Group and Roadmap WG to drive the vision and roadmap for 6G. It is a key contributor defining research areas, spectrum needs, use cases, and application requirements for the 2030 era.

At Nokia, we create technology that helps the world act together, and, at the same time, we act together with other stakeholders to define the next-generation communication system. We are thereby engaged in various collaborative research initiatives, alliances, industry collaborations, and pre-standardization bodies across the world, to drive our vision for a flexible, sustainable, trustworthy, agile and secure 6G architecture.

References

[1] Yannick Lair, “ 5G-Advanced’s system architecture begins taking shape at 3GPP”, September 2021, Online: https://www.nokia.com/blog/5g-advanceds-system-architecture-begins-taking-shape-at-3gpp/

[2] Nokia Bell Labs, “6G explained”, 2021, Online: https://www.nokia.com/about-us/newsroom/articles/6g-explained/

[3] V. Ziegler, H. Viswanathan, H. Flinck, M. Hoffmann, V. Räisänen and K. Hätönen, “6G Architecture to Connect the Worlds,” in IEEE Access, vol. 8, pp. 173508-173520, 2020, doi: 10.1109/ACCESS.2020.3025032

[4] China’s IMT-2030 (6G) Promotion Group, “White Paper on 6G Vision and Candidate Technologies”, June 2021, Online: http://www.caict.ac.cn/english/news/202106/P020210608349616163475.pdf

[5] Beyond 5G Promotion Consortium (B5GPC), Japan, https://b5g.jp/en/

[6] Korean Ministry of Science & ICT (MSIT), 6G, “Korea takes the lead once again ‘6G R&D implementation plan’ established”, Online: https://www.msit.go.kr/eng/bbs/view.do?sCode=eng&bbsSeqNo=42&nttSeqNo=517

[7] Petter Vetter, Nokia, “Global 6G Initiatives, Research Landscape”, July 2021, Online: https://www.wilsoncenter.org/sites/default/files/media/uploads/documents/07-08-2021%20-%20Vetter_6G%20Panel%20_2021-07-08_v04.pdf

[8] NGMN Alliance, “6G use cases and analysis”, February 2022, Online: https://www.ngmn.org/wp-content/uploads/220222-NGMN-6G-Use-Cases-and-Analysis-1.pdf

[9] Next G Alliance, https://nextgalliance.org/

[10] Next G Alliance, “Next G Alliance report: Roadmap to 6G”, Online: https://nextgalliance.org/white_papers/roadmap-to-6g/

[11] Hexa-X European 6G flagship, https://hexa-x.eu/

[12] Hexa-X Deliverable D1.5, “Initial 6G Architectural Components and Enablers”, Online: https://hexa-x.eu/wp-content/uploads/2022/01/Hexa-X_D5.1_full_version_v1.0.pdf

[14] Hexa-X Deliverable D1.2, “Expanded 6G vision, use cases and societal values”, Online: https://hexa-x.eu/wp-content/uploads/2021/05/Hexa-X-D1.2_v1.0.pdf

[13] Hexa-X Deliverable D1.3, “Targets and requirements for 6G - initial E2E architecture”, Online: https://hexa-x.eu/wp-content/uploads/2022/03/Hexa-X_D1.3.pdf

[15] Harish Viswanathan et al., “Joint Design of Communication and Sensing for Beyond 5G and 6G systems”, April 2021, Online: https://www.bell-labs.com/institute/white-papers/joint-design-communication-and-sensing-beyond-5g-and-6g-systems/

[16] United Nations, “Sustainable Development Goals”, Online: https://www.undp.org/sustainable-development-goals

[17] Gilberto Berardinelli et al., “Extreme communications in 6G: Vision and challenges for ‘in-X’ subnetworks”, November 2021, Online: https://www.bell-labs.com/institute/white-papers/extreme-communications-in-6g-vision-and-challenges-for-in-x-subnetworks/

[18] 3GPP, TS 23.288, “Architecture enhancements for 5G System (5GS) to support network data analytics services (Release 17)”, V17.3.0, December 2021,
Online: https://www.3gpp.org/ftp/Specs/archive/23_series/23.288/23288-h30.zip

[19] 3GPP, TS 28.533 “ Management and orchestration; Architecture framework “, V17.1.0, December 2021,
Online: https://www.3gpp.org/ftp/Specs/archive/28_series/28.533/28533-h10.zip

[20] Wikipedia, “Metaverse”, Online: https://en.wikipedia.org/wiki/Metaverse

About Nokia

At Nokia, we create technology that helps the world act together.

As a trusted partner for critical networks, we are committed to innovation and technology leadership across mobile, fixed and cloud networks. We create value with intellectual property and long-term research, led by the award-winning Nokia Bell Labs.

Adhering to the highest standards of integrity and security, we help build the capabilities needed for a more productive, sustainable and inclusive world.

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

Document code: CID212403 (April)