6G orchestration

A scalable approach to application-aware network optimization

White paper

This white paper focuses on principles of 6G orchestration and automation. It elaborates on the themes of the earlier Nokia 6G architecture white paper. We shall address drivers for 6G orchestration and automation (O&A), selected technologies relevant to the scope, and discuss the role of 6G O&A in monetizing service providers' capabilities.

Contents

Introduction

Orchestration is a technique that aligns a request with applications, data and infrastructure. Orchestration, in the context of cloud computing, provides a directed action towards goals and objectives having a wide scope [1]. In this white paper, we are specifically interested in network and service automation.

A number of current networking trends are notable:

- Communications service providers (CSPs) are deploying 5G networks
- Value networks associated with service provisioning are becoming more diverse
- Service providers are utilizing cloud platforms
- Enterprise-related connectivity and services are growing in importance
- Security and trust are of fundamental importance
- Application developers are increasingly important.

As a consequence of the above and other factors, the network automation and orchestration (NAO) technology field has several high-growth segments: 5G, software-defined networking, including software-defined wide area networking (SD-WAN), and multi-domain orchestration. The roll-out and optimization of 5G benefits greatly from automation, and consequently 5G NAO is forecast to achieve a 38% compound annual growth rate (CAGR) in the period 2021-27 [2].

General drivers for orchestration and automation (O&A) relate to simplification and reduction of operational expenses (OPEX). In evolving from 5G to 6G, automation will also enable use cases that would not be economically feasible without it. OPEX reduction techniques include domain automation and reduced configuration errors and rollbacks, for example, by means of intent-based networking and management. New use cases include co-existence of cloud radio access network (RAN) with classical 5G RAN, software-defined access networks, open RAN (O-RAN), end-to-end network slicing in 5G standalone (SA), multi-vendor network function (NF) related processes, process integration of development and operations (DevOps), and continuous integration / continuous deployment (CI/CD) pipelines. As we shall see below, the complexity of business and technical aspects of networks are expected to increase in evolving networks towards 6G.

The evolution of 5G to 6G brings new requirements and use cases for O&A:

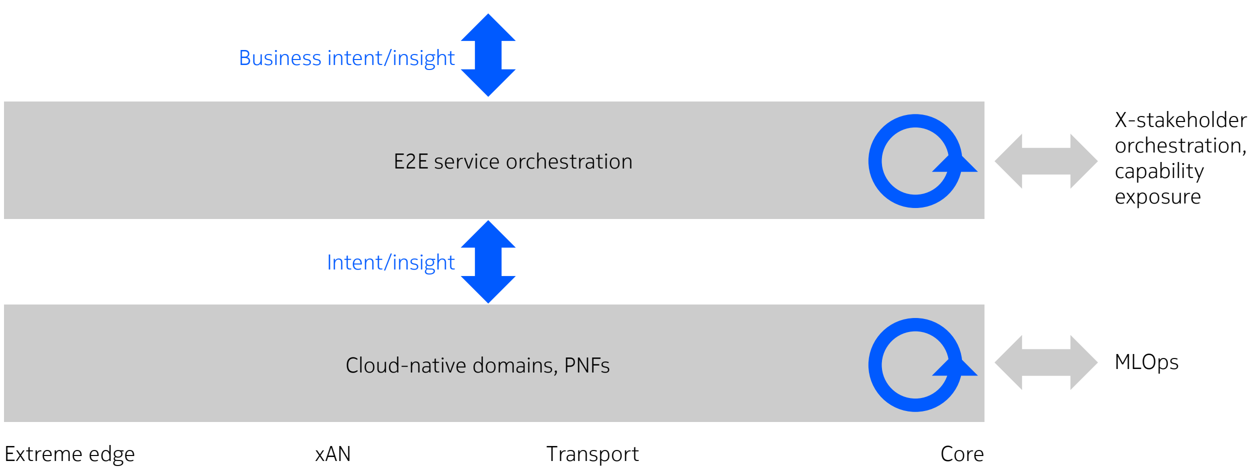
- Extended reality (XR) and the metaverse are expected to grow in importance and make use of advanced features of 5G networks
- CSP ecosystems will increasingly involve webscales
- The far edge of 5G RAN is expected to be an important part of the 6G infrastructure for most operators
- Network functions will evolve into cognitive functions (CF), which will necessitate orchestration of the AI/ML pipelines and related data
- Sustainability related orchestration will grow in importance
- Multi-access orchestration will be increasingly important in 6G
- Joint orchestration of connectivity and computing will increase efficiency and performance of the access networks
- Extending the scope of intent-based networking and management to RAN is expected to be important for automation of 5G-Advanced networks

- Interoperability of heterogeneous networks will be key for 6G
- Network exposure for applications (both user plane and network internal ones) will need to evolve for efficiency and performance reasons
- It is anticipated that there will be increased need for specialized non-public networks (NPNs) catering to the specialized needs of enterprise and industry segments.

Publicly funded projects are currently building consensus on the details of the next generation of networks. The most well-known is Hexa-X [3] and the recently announced Hexa-X-II follow-on project being funded by the European Union. The 6G network will be AI-native and have advanced analytics supporting various aspects of the system. New capabilities like joint communication and sensing (JCAS) are expected to be a part of the 6G system. The consequences of 6G principles like uniform orchestration interfaces towards cloud-native domains and AI-native design are being addressed by the work packages of Hexa-X.

Computing and network convergence (CNC) is being studied to create an architectural basis for unified orchestration of resources for 6G. 6G orchestration will need to support automated participation of multiple stakeholders to E2E service provisioning at all the stages and roles (resources or services). Nokia's views on general 6G architecture are summarized in [4]. In this white paper, we provide additional remarks relating to O&A domain. Figure 1 depicts selected key concepts of 6G O&A which will be discussed in subsequent sections referred to in the caption below.

Figure 1. Conceptual architecture of 6G network. Key components include intent-based networking/management (6G automation), controller architecture based on closed loops (6G automation), cloud-native technologies (technologies for the 6G edge), and capability exposure (service orchestration innovations). Machine learning operations automation is a cross-system enabler for 6G.



Based on the emerging consensus, some observations can be made about expected business aspects of 6G networks. New value networks will come into being in the context of use cases relating to human augmentation and digital-physical fusion. A greater diversity of service and technology providers are expected to participate in 6G end-to-end services. This development enables truly global service chains. Traditional providers can monetize their resources through a variety of capability exposure mechanisms. Applications interacting with the network will be more prominent in the design of 6G than in previous generations. Dynamic service instance composition is being studied with specialized capability exposure mechanisms that will allow traditional service providers to participate alongside natively distributed ones. Automated optimization of the network will be instrumental for 6G.

Some of the questions addressed in this white paper are:

- How are future networks managed?
- What is the cloud-native baseline for evolution from 5G to 6G?
- What are the technical enablers for 6G O&A?
- How are 6G capabilities monetized?

Below, we shall explore these questions through the following four focus topics:

1. 6G automation
2. Technologies for the 6G edge
3. Service orchestration innovations
4. Enterprise

6G automation

Monetization of CSP resources, in general, requires an ability to perform service innovation to provide for strategic differentiation in the marketplace, which contributes to top-line growth. Starting with 5G, network slices provide a platform for service innovation. As we shall see later, other service exposure methods provide alternative ways of leveraging value networks, thus monetizing CSP resources in novel ways during the evolution from 5G to 5GA to 6G. The ability to utilize a variety of value network interfacing mechanisms is important in view of the strategic differentiation within a marketplace as well as the quick evolution of the edge computing related business ecosystem.

Scalable network slicing in 5G requires automation of business, service and network management. In essence, advanced automation allows zero-cost customization of networks. Network owners can focus on the business opportunities their value network enables rather than having to deal with the complexity of the network itself. AI/ML capabilities are being added to the management layer as part of 5G standard evolution. Some of the trends discussed below extend the scope of technologies being developed during 5G evolution to meet 6G requirements.

The 6G network is expected to be more closely integrated with applications than 5G. To support this increased level of integration will require automated optimization of network configurations in response to variable loading patterns. For 6G, stricter performance requirements for certain traffic types imply that orchestration of transport jointly with services will be important. The 6G access network is expected to have more dynamic loading patterns than 5G, due to some of the network cloud platforms being closer to the network edge where the aggregation level is lower. Both trends will require advanced automation. Automation in 6G will support network programmability by end users, applications and operations personnel. Automation is also key in enabling operability for specialized NPNs through modularization.

In contrast to 5G, 6G will be AI/ML native, being defined for AI/ML from the ground up. For AI/ML to be used across the 6G system, however, the infrastructure needed by different variants of AI/ML must be mature, automated and support a wide palette of state-of-the-art techniques. This relates to data architecture, for example, as well as model management pipelines. Automated machine learning operations (MLOps) is a key enabler for 6G.

At the network edge, AI/ML is best viewed as a federated capability, including data collection for training and inference as well as bringing algorithms close to data. All aspects of the general AI/ML need to be scalable to support instantiation for various networking contexts, including the extreme edge. Data fusion will integrate data produced from diversified sources, e.g., sensors throughout smart cities, e-health sensors, and robots in advanced manufacturing.

Given the growing complexity of business and technological aspects of a 6G network, it is increasingly important to be able to consistently model the effect of changes to the network on service quality prior to committing the changes. A network digital twin supports this, enabling optimization of configurations prior to operationalization. Digital twins can be viewed as the networking equivalent of the digital-physical fusion of the metaverse [5]. Automated reasoning on operational triggers enables predictive orchestration for 6G.

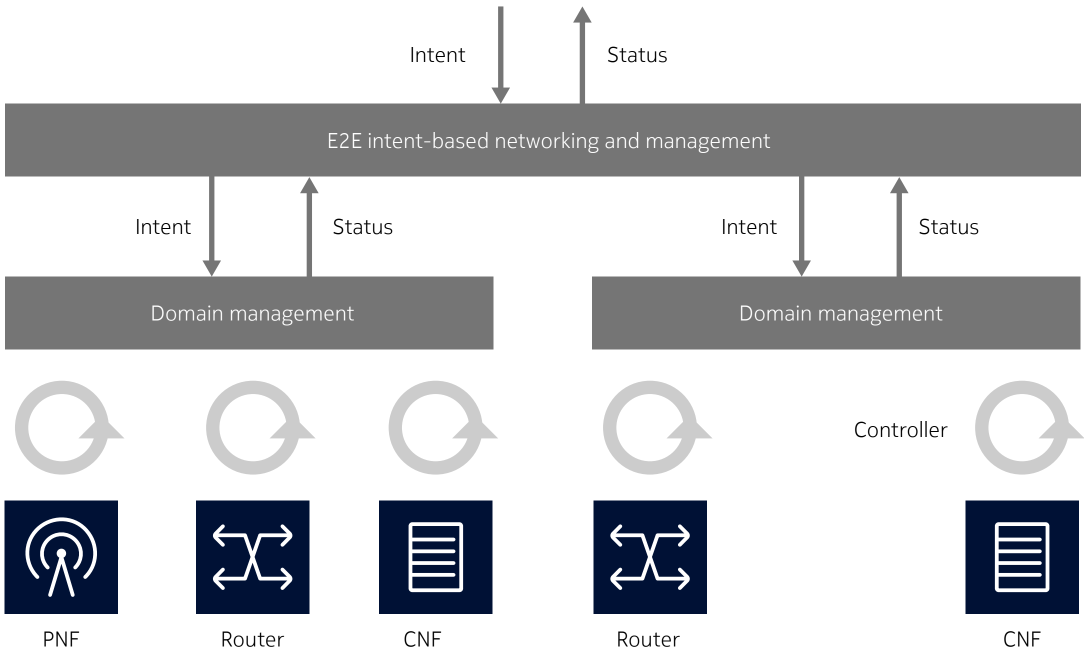
Automation is also reflected in how future networks are managed. Traditional network configuration has been imperative, amounting to step-by-step instructions for systems relevant to configuration changes. The state-of-the-art paradigm is declarative, where the goal state of the system is specified, and automation takes care of detailed steps needed to achieve the goal. This approach is used in intent-based networking and management, which is currently being extended from the transport domain to radio access networks. A business-intent interface provides an exposure mechanism for the network, enabling service instantiation and service assurance using domain concepts of the user.

The declarative paradigm is supported by a hierarchical automation architecture. Each layer has an incoming intent interface and outgoing feedback interface. The user of the intent interface (human or machine/application) should have the means to provide requirements freely (e.g., as text), which is mapped to relevant network capabilities by the automation layer to which the requirement is directed.

The network may engage in a dialogue with the user to fine-tune the requirements, aligning it not only with abstract network capabilities but also with the available resources. This helps the end user or operations staff to create an intent that is both ingestible and autonomously fulfillable by the network. Such intent-based operation requires that the network has reflection capabilities, in other words, it can model its components, interfaces and capabilities.

The modelling of the network needs to include both syntactical and semantical aspects to enable autonomous discovery of all interworking patterns between individual networking (NF and management) components. Semantical compatibility is especially important to ensure that two components orchestrated to interface with each other create/consume information conforming to the same semantics. Intents are relevant to the scope of the layer in question. The E2E service management layer is driven by business intents, whereas layers lower down in the hierarchy receive service, technology domain or resource facing intents. Figure 2 illustrates an automation architecture based on intents.

Figure 2. Hierarchical intent-based management architecture. For simplicity, only two layers of an intent-based management architecture are depicted. PNF signifies physical network function and CNF cognitive (virtual) network function.



Within each layer, the incoming intent is translated to technical parameters relevant for the layer in question. AI/ML technologies are used to create predictive models of the network, including state models for quantifying the behavior of network functions, services, devices and users, topology and resource models. In addition, there are semantic models that define the capabilities and interfaces of each network and management component. These models enable detection and prediction of conditions that require automated resource or service management actions.

The outcome of the mapping is a technical goal state. ML-enabled closed loop automation (CLA) algorithms, exemplified by self-organizing network (SON) use cases for 5G RAN, take care of adjusting the technical parameters to achieve the goal state. In essence, a closed-loop algorithm implements a dedicated controller that is configured by means of intents. Combined with AI/ML, 6G CLA-based controller architecture is a key building block for a 6G cognitive autonomous network (CAN) [6].

Zero-touch automation approaches to network configuration, orchestration, operation and service chaining are being developed but more work is needed for E2E orchestration given the increased heterogeneity of domains expected to be relevant to 6G. It is expected that orchestration in 6G will be more complex due to introduction of new criteria like sustainability. Cross-domain automation is the key to unlocking the potential of a fully automated network that can deliver end-to-end network slices on demand over multiple domains of one service provider as well as for services involving several service providers. Intent-based networking is thus best driven by E2E intents, which are then refined into domain specific intents. Unified and streamlined data and information models and standard command protocols are important for coordination of business processes and network service lifecycles across stakeholders.

Ultimately, network programmability is brought by design to the 6G ecosystem, masking the ecosystem complexity from service providers, verticals and enterprises. Network programmability allows for instantiation and operation of services involving components in different domains and layers, building on configuration of resources in the respective domains.

Technologies for the 6G edge

The network edge is the meeting point of stakeholders as well as technology domains. Consequently, the edge is the focus of rapid technological innovation at the moment. Both CSPs and enterprises are exploring options for offloading some their workloads to cloud computing service providers. The evolution of cloud platforms proceeds in 5G from core network clouds towards regional and further to far edge clouds. In enterprises, the evolution proceeds from the cloud providers' big data centers, on the one hand, and on-premises computing platforms, on the other, to low-latency off-premises edge clouds. Neutral hosts are growing in importance in hosting technological capabilities for CSPs.

A future-proof cloud platform for both enterprises and different types of service providers supports management of workloads from multiple providers. Security and privacy are of utmost importance for edge cloud deployments and require architectural solutions. For a 6G multi-access network to evolve to meet the new requirements, the slicing paradigm will need to be extended to include novel 6G aspects.

Evolution from 5G to 6G will further accentuate the network edge. Not only will cloudified radio access network functionalities be located closer to radio heads, but draft orchestration architectures considered in the Hexa-X project also allow orchestration of the "extreme edge" beyond the mobile network as part of end-to-end services. Examples of extreme edge include enterprises and home networks. The extreme edge networks can function as standalone networks with dedicated orchestration functionality or be orchestrated as a part of E2E 6G services spanning 6G CSP networks. The orchestration of the extreme edge is supported by the AI/ML federation capabilities mentioned in the previous section.

A candidate 6G principle from the Hexa-X project is continuum orchestration across the capabilities of the 6G network. Orchestration of far edge and extreme edge resources needs to ensure adequate performance levels. In essence, this means moving from domain-specific orchestration interfaces to generic ones that can be utilized all the way from the network core to the extreme edge. The 6G extreme edge, in general, includes systems beyond the far edge of the network, including user equipment (UE) as well as on-premises resources that can be orchestrated as part of E2E 6G services. On-premises resources can also act as an autonomous subnetwork when not part of an E2E 6G service.

The cloud-native paradigm employed in the domains needs to be compatible with the intent-based management/controller paradigm, and it must scale for different aggregation levels ranging from network core to enterprise and home domain networks. Nokia is participating in the Nephio project [7], which studies the use of Kubernetes-based technologies for cloud-native orchestration. The 6G-specific aspect concerns how cloud-native will be used in the 6G access network architecture and supporting 6G value networks.

An important aspect of the cloud-native paradigm is coordinated management of processes enabling true multi-vendor cognitive functions to be used in the domains. This includes both lifecycle management aspects of individual CFs as well as coordinated CI/CD pipelines across the CFs.

The network edge needs to cater to the federated AI/ML mentioned earlier. Depending on the approach, data collection for training and inference needs to be in place and execution platforms for offloading of inference need to exist. Federated learning is a candidate technology for dealing with the context specificity of ML models of limited size at the network edge. AI/ML interworking between domains is expected to be important for some use cases. Due to the low aggregation level of edge environments, the provenance of context information associated with AI/ML models is crucial.

Service orchestration innovations

Service orchestration innovations complement the technologies described above by supporting a variety of 6G business scenarios involving different constellations of stakeholders. Run-time multi-stakeholder service design is expected to be important for supporting a diverse provider ecosystem for 6G. As mentioned earlier, developers and system integrators are increasingly important for service delivery.

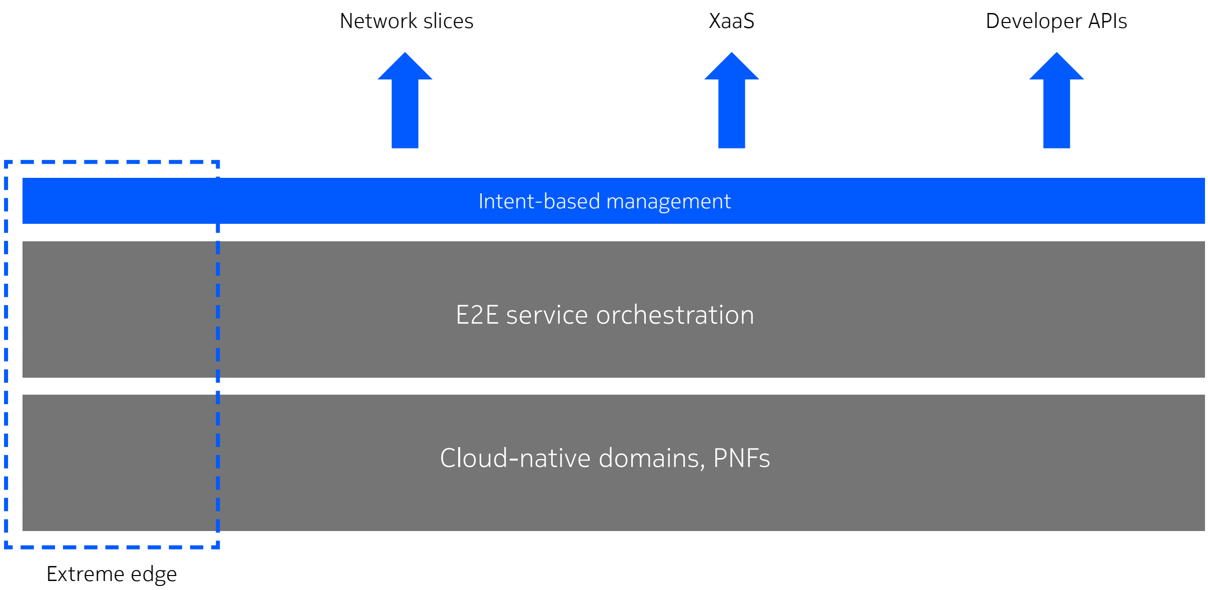
Network slices are a powerful channel for monetizing CSP resources. It is expected that network slice lifecycles are more dynamic, have finer granularity for service provisioning, and can be geographically more localized in 6G than 5G. Network slices extending to the extreme edge will require new technological solutions.

Continuum orchestration including extreme edge, subnets, intent-based management, and AI/ML federation are expected to be relevant to enterprises in a 6G world, and indeed are candidate technologies for 6G orchestration. Joint orchestration of connectivity and computing resources is expected to be important for the edge of the network. Sustainability related orchestration is another opportunity for 6G, for instance, opportunistic use of localized/surrounding resources (network, IT, devices) to optimize energy consumption and sharing. It is also conceivable that the scope of 6G end-to-end orchestration will be extended to new areas such as radio spectrum.

In view of the variety of needs of enterprises, it is also relevant to consider network capability exposure mechanisms beyond network slices to monetize CSP resources or integrate NPNs into production environments. These include programmatic exposure of network capabilities via application programming interfaces (APIs) targeted towards developers building bespoke integrations for enterprises or to enable everything-as-a-service (XaaS) developers to automate service-level integration for other providers.

The API approach could, for instance, allow for the creation of something called the 'eventverse', where participants have powerful capabilities to personalize their experiences, such as sporting or cultural events. XaaS programmatic exposure might enable distributed orchestration of an end-to-end service combining capabilities of participating domains in run time. The domains could include CSPs, IT providers, or specialized application providers for services such as extended reality (XR). Dynamic registry and discovery mechanisms for heterogeneous and highly distributed services (network, data, enterprises, application) are some of the key assets for building a unified service mesh. The goal is to enable flexible composition of capabilities and services (see Figure 3).

Figure 3. An illustration of service exposure from a 6G CSP network. Intent-based management supports different exposure methods.



Security and privacy design is a primary concern when orchestrating services from third-party providers and requires the establishment of liability sharing control mechanisms based on, for example, delegation models and policy-based control. In an extreme case, service composition is performed dynamically in fully peer-to-peer (P2P) fashion without assuming a shared trust between providers. In such a case, distributed ledger technologies (DLTs) have high potential. Capability marketplaces form a middle ground between exposure of selected CSP capabilities and P2P orchestration.

Industrial use cases

Industrial use cases illustrate the technologies and principles introduced in the preceding parts of this white paper. They also provide a context allowing us to introduce some new concepts.

By the dawn of the 6G era, many industries are expected to have integrated 5G network technologies as part of their production environment. Catering for the long tail of enterprises requires networks to be progressively easier to integrate with production environments and operate. For industrial applications, the 5G network provides predictable service quality for the different use cases relevant to the production environment. The 5G architecture also provides edge computing capabilities as an integral part of the communications architecture. The hallmark of 5G for industrial applications is high reliability. For industrial applications, support for developers and system integrators is also important.

Consequently, as 6G launches, the adoption of advanced connectivity by industries is expected to have grown and individual production environments become technologically more advanced. Technological advances in production environments will include simple digital/physical fusion such as XR as well as more comprehensive application of digital twin and metaverse approaches. In short, as the range of industrial use cases becomes more varied, the value network joining 6G connectivity and computing providers will become more complex.

For industrial use cases of the future, the ability to federate resources flexibly and dependably is important. Each company's computing and communication resources could simultaneously act as a stand-alone orchestrated network or be orchestrated as part of 6G E2E services provided by CSPs or other providers. 6G will, thus, need a palette of technologies to support this expanded range of industrial applications. NPNs will support specialization to meet the specific needs of different industry segments [8]. CSP exposure mechanisms discussed in the previous section provide another way of achieving this. Support for application developers and system integrators is crucial to enable the full range of industrial use cases. It will be increasingly important for networking technologies to support connectivity and computing coordination via paradigms such as CNC mentioned earlier. Application exposure needs to evolve from NEF of 5G to more symmetric interaction in 6G with anticipatory interaction between the application and the network.

Security and privacy are important for industrial use cases and consequently an integral part of the emerging 6G system design. Many of the industrial use cases are expected to involve distribution of AI/ML-related processing. The AI/ML federation capabilities described above should support distribution of processing while ensuring security related aspects for data.

Perspective

The 6G value network of service providers and customers will be more varied than for 5G, which will be reflected in the diversity of use cases. Consequently, deployment scenarios are also expected to be more diversified. The 3GPP standard for 6G will provide a global baseline for implementations across a variety of use cases, blending in other specifications such as the O-RAN domain. Cloud-native relevant standards, de facto standards, and best current practices will complement and contribute to the 3GPP standard for specific areas like cloud-native computing. The interaction between standards developing organizations (SDOs), best current practice (BCP) developing organizations, and de facto standards is expected to become more important.

6G networks need to be designed with the 5G/6G transition in mind to ensure economic viability. The technical complexity of 6G is expected to be higher than 5G due to support for a wide variety of use cases and architectural changes such as the extended scope of the service-based architecture (SBA). As a result of changes in the radio access architecture, it is expected that the loading patterns at the network edge will be more dynamic in 6G as compared with 5G. Co-orchestration of applications, cognitive functions and relevant platforms and coordination of multi-vendor CF lifecycles will add to the complexity.

The evolution of industrial use cases during the 5G era will provide ongoing guidance as we approach the launch of 6G. Nonetheless, it is already clear that the long tail of enterprises and industries will benefit from automation of 6G networks, allowing both service providers and private networks to tailor specific services to meet varying needs and business models. CSPs will be able to use 6G to create extended value chains using exposure capabilities to foster co-creation of value with a network of partners. Thus, 6G will need to be programmable by application and end user — as well as the operational staff of service providers and third-party partners. The 6G ecosystem will need to support a heterogenous “network of networks” approach to integrate various stakeholders, which will require multi-stakeholder and cross-stakeholder orchestration. Automation will, thus, be essential to manage the increased complexity for operational staff, with declarative interfaces for intent-based networking proving critical for orchestration and administration (O&A).[1]

A 6G ecosystem building on 5G evolution provides a stable platform for the transition to 6G. This approach expedites the adoption of 6G and reduces economic risk for stakeholders. Overall, automation will have a more pronounced role in 6G as compared with 5G, supporting the strategic network monetization goals. Advanced automation of 6G networks will allow service providers to focus on value creation, fostering innovation and enabling new business models.

1 Note that as the degree of automation increases, the explainability aspect of automation will also rise in importance.

Abbreviations

AI/ML	Artificial Intelligence / Machine Learning
API	Application Programming Interface
BCP	Best Current Practice
CAGR	Compound Annual Growth Rate
CAN	Cognitive Autonomous Network
CF	Cognitive Function
CI/CD	Continuous Integration/Continuous Deployment
CLA	Closed-Loop Automation
CNC	Computing and Network Convergence
CSP	Communications Service Provider
DevOps	Development and Operations
DLT	Distributed Ledger Technology
IT	Information Technology
JCAS	Joint Communication and Sensing
MLOps	Machine Learning Operations
NAO	Network Automation and Orchestration
NEF	Network Exposure Function
NF	Network Function
NPN	Non-Public Network
O-RAN	Open RAN
O&A	Orchestration and Automation
P2P	Peer-to-Peer
RAN	Radio Access Network
SBA	Service-Based Architecture
SDO	Standards Developing Organization
SD-WAN	Software-Defined Wide Area Network
SON	Self-Organizing Network
XaaS	Everything as a Service
xAN	Multi-access Network
XR	Extended Reality

References

1. https://en.wikipedia.org/wiki/Orchestration_(computing).
2. Analysys Mason, Network automation and orchestration, worldwide forecast 2022-2027.
3. https://www.bell-labs.com/institute/white-papers/technology-innovations-for-6g-system-architecture/
4. https://hexa-x.eu/
5. https://www.nokia.com/blog/the-metaverse-will-never-move-beyond-our-living-rooms-without-a-powerful-network/
6. S. Mwanje and C. Mannweiler (editors), Towards cognitive autonomous networks: network management automation for 5G and beyond, John Wiley & Sons, 2020.
7. https://nephio.org/
8. https://www.bell-labs.com/research-innovation/what-is-6g/6g-technologies/

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

Document code: CID213047 (February)