Autonomous services and unified networking experience technology (UNEXT)

UNEXT whitepaper series

Vilho Räisänen

Automation is becoming increasingly important as the complexity of communication networks continues to grow and serve an expanding range of value networks. These developments are driving the trend toward greater autonomy of networks, powered by evolving AI/ML paradigms. Autonomous services are a fundamental building block for advanced network autonomy, whether in traditional hierarchical or emerging decentralized architectures. In this white paper, we describe the unified network experience technology (UNEXT) approach.

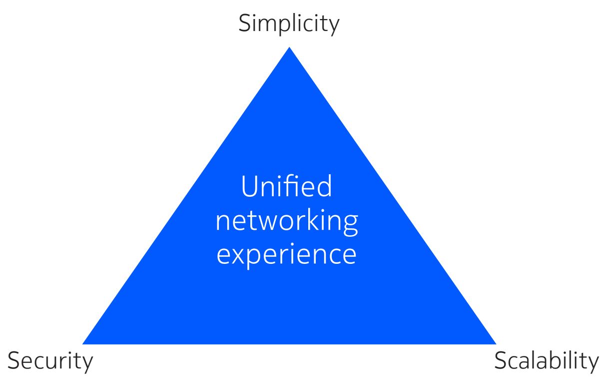
Contents

The UNEXT vision

As envisioned, unified networking experience technology, or UNEXT (Sefidcon, Vulkán, & al, 2023), is an intelligent networking system providing communications and computing in a simple, secure and scalable manner (Figure 1). It will enable networking to be a self-managing, interactive system that supports users and applications with resources and services to scale across heterogenous users, operators and execution environments.

Achieving these aspirations requires tackling the reasons why networking today is neither simple, secure nor easily scalable. In consideration of technology trends, five sets of capabilities were identified to confront and develop solutions for undesirable networking characteristics and behaviors. Once created, these technologies will simultaneously mitigate the causes of network complexity, insecurity and scalability impediments. The sets of capabilities defined for investigation are the following (Figure 2):

Figure 1. UNEXT Vision

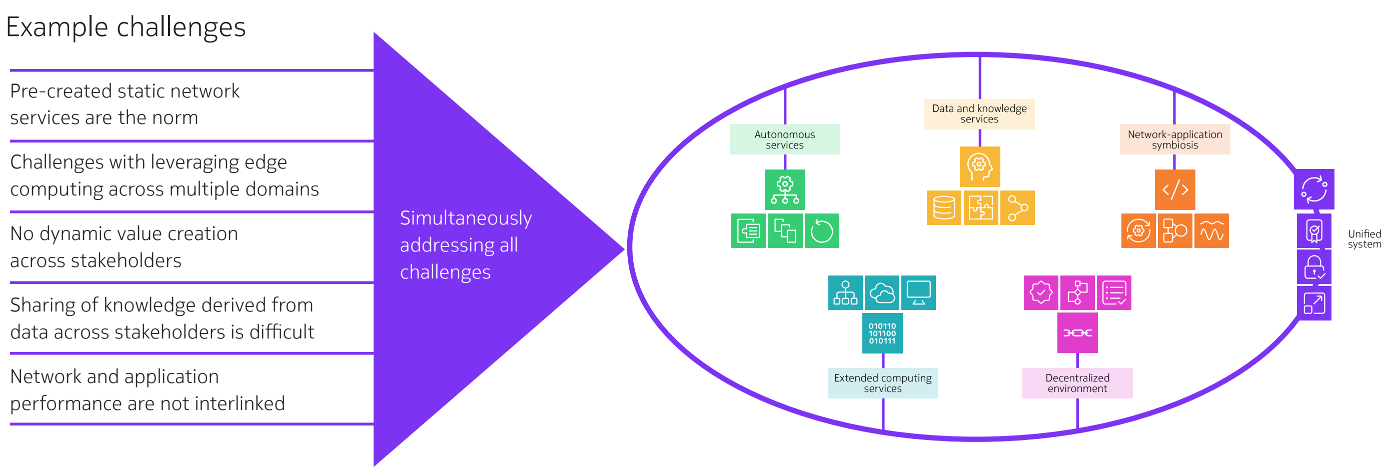


- **Autonomous services:** As networking systems become increasingly complex, the degree of human-guided automation is increasingly mismatched with the complexity of network operations and procedures needed to deliver services. Furthermore, service lifecycle operations such as service creation, provisioning, assurance, updating, maintenance and decommissioning are expected to scale beyond what human-involved network operations can deliver in terms of service complexity and request-to-delivery response times. Thus, autonomous decision-making and realization of services, including delivering, discovery, design, composition, monitoring and exposure of new services will be an essential capability in UNEXT.
- **Network-application symbiosis:** An important part of the networking experience is the combination of network services and endpoint applications that rely on them. User experience will be enhanced by enabling applications and the network to interwork and co-optimize carrier-grade application reliability, network usage and quality of experience.
- **Decentralized environments:** To provide end-to-end services, a UNEXT instance is required to interact with other domains, systems and resources under the control of their stakeholders. Services will still require completion even if there is no single centralized authority coordinating them, such as in the network cloudification trend with network functions and microservices running in a compute continuum. Thus, an essential capability will be enabling among heterogeneous stakeholders the end-to-end autonomous service coordination, including negotiation, data sharing, orchestration and assurance, all based on

dynamic determinations of trust.

- **Extended computing services:** Compute resources abound in networks that can be utilized for end-to-end service delivery and achieving sought-after quality of experience. To address the challenges and leverage these resources across the compute continuum, UNEXT defines the abstraction, orchestration and support for compute capabilities across the device-edge-cloud continuum.
- **Knowledge and data services:** At the core of the UNEXT system is the ability to accommodate the flood of data from all areas—users, operators, systems and environments—and to retain and update it as evolving knowledge, all in order to enable sound decision-making. This provides the reasoning scaffold for future data/AI-driven networking solutions.
- Importantly, the capability sets will overlap in terms of the technologies required to provide or address their associated use case categories. Furthermore, a collection of technologies alone is not sufficient for UNEXT. To develop this system design and technology integration, we need a unified system that ties the technical aspects together. The result will be an autonomous networking system that enables cross-stakeholder service provisioning and execution in a simple, secure and scalable manner by design.

Figure 2. UNEXT capabilities



Challenges and approaches

In this section, we discuss the challenges in service lifecycle management and provide a vision of how UNEXT can help to address them. We employ the concept of autonomy to introduce the UNEXT approach as a key building block to address topical industry challenges.

Challenges

Provision of services is a primary value creation element. A common challenge in O&A underlying service provision is monetizing a service provider's resources in a competitive environment in which differentiation requires agility in terms of the service portfolio. Value networks are evolving rapidly, particularly at the network's edge. What is needed is the ability to flexibly create new value network constellations, optimally utilizing existing capacity in each situation.

The legacy world of networks was based on imperative management, with each configuration change necessitating step-by-step instructions for each of the resources involved. The proliferation of proprietary interfaces to systems led to a jumble of bespoke integrations, which was in stark opposition to the goal of business agility.

Automation has also been pursued in networks through the application of AI/ML to a wide range of use cases in networking. Unfortunately, these implementations have typically been customized, leading to a multitude of implementations that utilize a variety of techniques. The complexity that the AI/ML was supposed to manage has been transferred to the variety of automation platforms involved. Additionally, the AI/ML algorithms need to be managed.

There's a beneficial set of techniques at hand to pave the way towards the future. What is needed is a vision to guide the evolution towards 6G, supporting monetization during the transition towards the future while also facilitating a spectrum of innovative business models. At the same time, automation needs to be addressed on an architectural level.

Autonomy as an enabler for agility

The principle of autonomy holds the promise of taking automation to the next level and enabling a level of business agility that the rapidly evolving business environment calls for.

Horn, Paul (2001) introduced the concept of autonomous computing, the provision of a capability over a service interface while exhibiting autonomous behavior and awareness of the relevant environment. The so-called self-X capabilities (e.g., self-configuration, self-optimization) are well-known aspects of autonomy. The autonomous computing concept inspired communication network automation paradigms such as self-organized networks (SON) (NGMN, 2007). More recently, autonomy at the network architecture level has been addressed by TM Forum (TMForum, 2022). The next challenge is evolving the industry towards higher levels of autonomy involving automation of service operations. In the TM Forum classification, AI is fundamental for enabling autonomy levels 4 and 5.

Some of the key aspects of autonomous network architecture include intent-based management and closed loops on different levels of the service and network management architecture. Recently, the importance of configuring assurance as part of service creation has been recognized. The possibility of creating solutions—potentially involving code generation—in real time using GenAI has raised a lot of interest.

Bell Labs vision for the evolution of orchestration and automation (O&A) towards 6G was outlined in (Räisänen & al, 2023). In addition to intent-based management architecture based on closed loops, it also featured the concept of network digital twin (NDT) distributed into the intent management hierarchy.

UNEXT and autonomy

One of UNEXT's central theses is that system software plays a critical role in evolving networks toward the future (Räisänen & al., 2024). Not only does the adoption of proper principles facilitate the autonomous transformation of traditional service provider networks, but it also enables novel, decentralized and no/low-trust value service provision value network constellations. Furthermore, current service providers can experiment with novel value networks alongside their existing partner setups.

In the previous sections, we have described how the capabilities of the UNEXT autonomous service vision support the transition of the telecommunications industry towards advanced automation solutions. A run-time composable autonomous service instance supports the declarative management paradigm and enables the use of rapidly evolving AI/ML techniques in the process. Intents provide an abstraction level for declarative management which, provides several advantages:

- Allow business interactions between stakeholders using concepts familiar to them.
- Facilitate closer collaboration between applications and the network.
- Tailor interactions to the context.

Support for decentralization is a hallmark of UNEXT. The autonomous services can leverage it for discovery of service capabilities in other UNEXT domains. In fully decentralized scenarios, this can be performed in a no/low-prior trust setting.

The extended computing services approach of the UNEXT vision takes care of instantiating the service blueprint in specific execution environments and orchestrating requisite services, including data, AI/ML models, and communications. A particular service instance might involve multiple administrative domains and types of execution platforms within each domain, ranging from general central processing units (CPUs) to task-specific accelerators.

In addition, the UNEXT architecture vision supports autonomous services by providing knowledge and data services that complement the service composition. Service blueprints specify the types of data services and AI/ML models needed by the service instances, while UNEXT architecture capabilities ensure the provision of the appropriate instances relevant to the deployment target. For instance, a controller of a particular component of a base station would be linked to the data feeds relevant to the unit. Similarly, the service instance would have access to AI/ML models trained with the respective data, which, if not readily available, would trigger MLOps training.

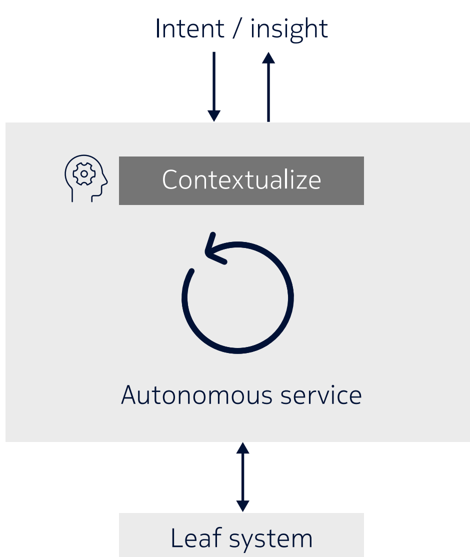
Autonomous service

In this section, we describe autonomous service as a fundamental architectural unit that uses UNEXT principles (Sefidcon, Vulkán, & al, 2023) and capabilities to provide autonomy for both hierarchical and decentralized service provision scenarios. Furthermore, we describe a method for composing services in run time.

Autonomous service

The Nokia Bell Labs vision for an autonomous service (AS) instance is illustrated in Figure 3. It includes an intent/insight interface, contextualize functionality, autonomy mechanisms and an optional proprietary interface towards leaf systems. An AS instance may consist of a single service or be composed of multiple services, which together constitute the autonomous capability. In the UNEXT approach, service composition takes place when there is no function available to handle an incoming intent. We shall describe the composition aspect later in this paper.

Figure 3: Autonomous service instance. The interface towards leaf systems is relevant for controllers of resources and is thus optional.



An insight is a return channel for intents on the same conceptual level. An insight differs from a traditional intent report in that it is expressed with the same domain concepts as the intent. Thus, for example, an AS instance communicating with a human user via a natural language interface would also issue natural language insights as responses to the intents.

The contextualize functionality implements the environmental awareness aspect of an AS instance and is used to interpret intents and create insights. The contextualize functionality involves analytics and external links necessary for this. In accordance with the distributed NDT principle, we consider the contextualize functionality to include a NDT tailored to the scope of the AS instance. We shall describe the composition of the NDTs later in this paper. The UNEXT system will provide knowledge and data services for contextualization, including:

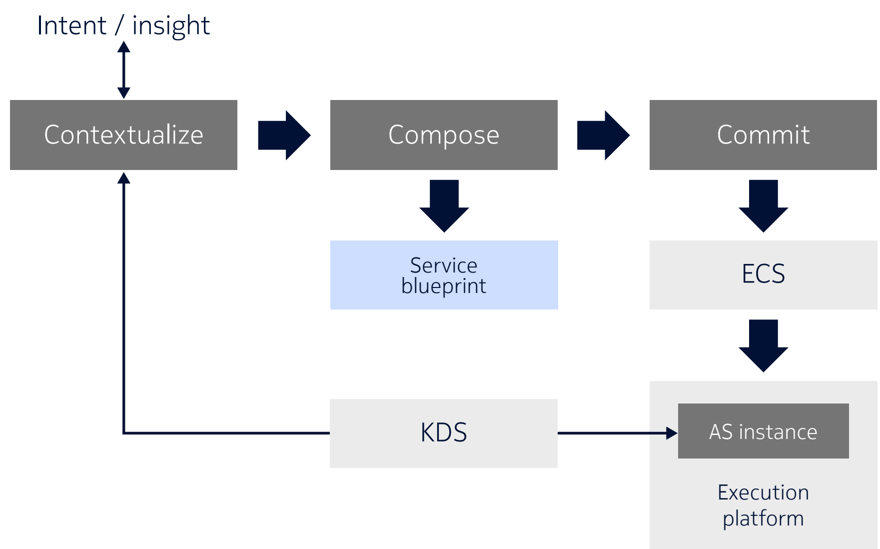
- Data and models used by the NDT
- GenAI support for the creation of insights

The autonomy mechanisms implement the self-X aspects of an AS instance and may involve closed-loop automation (CLA) alongside a variety of potential artificial intelligence/machine learning (AI/ML) methods. Thus, for example, potential GenAI capabilities should be embedded within an atomic AS instance or a constituent service for a composite AS instance. The UNEXT system will provide AI/ML model instances relevant to the execution environment, triggering training of new models with AIOps as necessary.

Service composition

The autonomous service concept of UNEXT includes the composition of services in run time, triggered by an incoming intent (Figure 4). No prior design phase is involved in service composition.

Figure 4: Illustration of service composition.



Service composition involves a contextualization phase for the incoming intent analogous to the functioning of an AS instance.

Service composition uses information about existing services in the same autonomous domain as well as other domains, the latter supported by UNEXT decentralization mechanisms. The Natural Language Network demonstration of Nokia (Nokia, 2023) is based on semantic modeling of service application programming interfaces (APIs). Service composition includes the assurance capabilities of the composed service as well as AI/ML capabilities potentially needed. The output of the service composition phase is a service blueprint.

The third phase of autonomous services architecture is commit, which deploys the service instance in the cloud/compute continuum with the necessary connectivity and orchestrates the necessary data/knowledge service instances. Some of the components may already be running and are linked to the composition. Commit makes use of the extended computing services capability of UNEXT.

Once the components of the service instance have been deployed on one or more execution platforms using UNEXT extended computing services, the originator of the composition intent is informed, and the intent/insight of the AS instance is activated. The originator of the intent is appraised of the status of the newly created service instance continuously by means of insights, providing a more advanced level of feedback as compared to traditional intent reports.

Perspective

In this section, we put the autonomous services concept into perspective through industry frameworks and potential use cases. The first example links back to the concepts of Section 2, elaborating link to autonomic networks. The second example relates to business agility and service creation. The third one, in turn, shows that the principles of service composition can be applied to digital twins more widely than just in network context.

Autonomous network

The autonomous network architecture of TM Forum (TMForum, 2022) includes the Intent Handling Function (IHF) as a basic building block. The AS vision for autonomous service of Bell Labs shares capabilities with IHF, including intent interface and self-X aspects. The tailored NDT inherent in AS instances goes beyond the IHF, as it provides tailorability to different use cases.

The service composition capability of autonomous services allows for a flexible creation of autonomous services on demand, in contrast to the typical service creation process involving a separate design phase. In the UNEXT approach, assurance is an integral part of the service composition. The composition approach allows for the use of state-of-the-art AI/ML methods for autonomy and avoids being limited to one paradigm only.

The UNEXT architecture brings with it the data and knowledge services that can be used by the NDT and insight generation. Furthermore, the data services and locally relevant AI/ML models provided by the UNEXT architecture support the autonomy of the AS instance.

Business agility for services

Since the intent/insight interface is not limited to specific intent formalisms, customer-facing services may use natural language intents and provide feedback about service performance through insights in the same manner. For the latter, LLM models can be leveraged.

Semantic modeling of the APIs of resource-facing services makes it possible to rapidly create top-down services. For services interfacing to the resources (e.g., controllers of physical or virtual systems), the intent/insight interfaces are automatically adapted to the adequate abstraction levels derived from the semantically generated API descriptions.

Digital twin composition

As a part of the overall composition of the AS instance blueprint, a tailored NDT relevant to the scope of the AS instance can be created. This involves linking in the relevant models and knowledge services.

More generally, any digital twin itself can be viewed as an autonomous service and be composed in the same way as other services. Consequently, the autonomous services approach—supported by knowledge and data services—enable the creation of a network of digital twins across communication networks and application domains. Specifically, tailored digital twins can be created on demand to serve a transient need with specific requirements.

Summary

UNEXT's autonomous service vision of UNEXT provides capabilities that are aligned with autonomous networks based on intent interfaces. The concept of autonomous service is general, allowing incorporation of rapidly evolving AI/ML technologies into automation of the network. The ability to compose services in run time without a separate design phase is supported by automated service modeling capabilities.

Together with other UNEXT working areas, they support the transition from current-day systems towards effective alignment of networks with market needs while allowing for effective use of resources. In addition to traditional hierarchical O&A systems, the concept of autonomous services is compatible with fully decentralized, no/low trust networks.

References

Horn, Paul. (2001). Autonomic computing vision. IBM.

NGMN. (2007). Next Generation Mobile Networks Use Cases related to Self Organising Network, Overall Description. NGMN.

Nokia. (2023). Nokia unveils AI breakthrough that can reconfigure networks through human speech. Retrieved from Nokia: Nokia unveils AI breakthrough that can reconfigure networks through human speech.

Räisänen, V. (., & al. (2023). 6G orchestration: a scalable approach to application-aware network optimization. Nokia Bell Labs.

Räisänen, V., & al., e. (2024). 6G orchestration and automation: a system software view. Nokia.

Sefidcon, A., Vulkán, C., & al. (2023). UNEXT: A unified Networking Experience. Bell Labs.

TMForum. (2022). Autonomous networks technical architecture, IG 1230. TM Forum.

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

Document code: CID214368 (January)