UNEXT:
Network-application symbiosis

UNEXT White Paper Series

Sarit Mukherjee, Andrea Francini, Michael Montag

Contents

Executive summary

Network-application symbiosis is one of the key capabilities of the Unified Networking Experience Technology (UNEXT) [1] vision. Its realization requires the exploration and development of technologies that enable immediate accommodation of the application session and preserve the satisfaction of its requirements through all contextual transitions. All symbiotic solutions pair the distributed realization of their functional capabilities with their centralized exposure. As a result of the symbiotic relationship, the UNEXT system provides every application with a run-time platform that has a dedicated, custom-tailored composition of compute and network capabilities. The application, in turn, generates new revenue opportunities for the UNEXT provider.

Network-application symbiosis: A new trend for network applications

Until today, applications and networks have been operating in a general state of reciprocal isolation. As a result, they have been engaged in a typical arms race: as networks become more capable, both functionally and resource-wise, applications grow more complex, dynamic and demanding. While every cycle of the race enriches the application landscape with new opportunities and sources of revenue, network service providers keep struggling to extract their fair share of the bounty.

A novel technological trend, which we call network-application symbiosis, is now creating a wide opening for a friendlier relationship between applications and networks: emerging applications, across consumer, enterprise and vertical markets, are increasingly dependent on security, high reliability and low latency, to a level that the over-the-top (OTT) model for application deployment and operation can no longer sustain. These new mission-critical applications, such as autonomous driving, industrial control and the metaverse, cannot operate in an environment of obliviousness between application and network; they can only thrive on a foundation of reciprocal awareness and cooperation.

The UNEXT network-application symbiosis principles encompass the mutualistic interfaces, functions and architectures that will deliver the full potential of the new applications and untap steady sources of revenue for the networks that support them. Their successful realization will be measured for every application session by the uninterrupted fulfillment of its intents and by the value that it brings to every network segment over which it extends.

The network stripe constructor

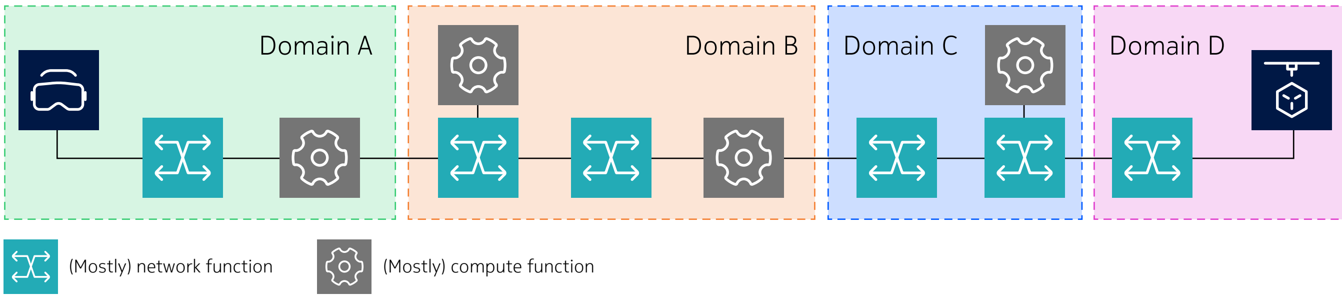
On the network side, uninterrupted intent fulfillment for every application session requires knowledge of the session context and the ability to immediately adapt to its evolution, together with fine-grained control of the network-compute resources that the session is allowed to consume. At the same time, the application must maintain continuous visibility over the quality it is achieving, using key quality indicators (KQIs) gathered by its endpoints and key performance indicators (KPIs) measured by the network, and then turn observations into corrective actions whenever the inference from the stream of indicators deems it necessary. Rapid reciprocal adaptation means that application performance and user satisfaction can be improved with fewer (incremental) resources, thereby providing the network with the benefit of higher revenue at a lower cost.

Figure 1. The three pillars of network-application symbiosis: mutual context awareness, dynamicity, and granularity.

As shown in Figure 1, mutual context awareness, dynamicity, and granularity are the critical dimensions for the realization of the symbiosis. **Mutual context awareness** results from continuous exchanges of information. The frequency and volume of the exchanges depend on the transparency of the interface between application and network, which scales with the established level of reciprocal trust. On each side, **dynamicity** quantifies the promptness of reaction to contextual changes manifested by the symbiotic counterpart, starting at the time of instantiation of the application session. **Granularity** refers to the breadth of the information that the two sides can exchange and to the level of functional differentiation that the network can provide, first across applications and then between sessions of the same application.

Upon instantiation of a new application session, UNEXT establishes a network path (see Figure 2) that not only interconnects all the session endpoints but also provides all the functions and service features that fulfill the intents of the application. The result is a dynamically composed and controlled chain of compute and network functions that presents to the application the behavior of a fully dedicated data and services path with all the required network-compute resources. For ease of discussion, we call this functional chain a network stripe, emphasizing that the term only refers to the behavior perceived by the application and does not mandate specific architectures, functions, or protocols for its realization.

Figure 2. A network stripe is a dynamically composed chain of compute and network functions spanning across multiple domains that presents to the application the behavior of a fully dedicated service and data path.



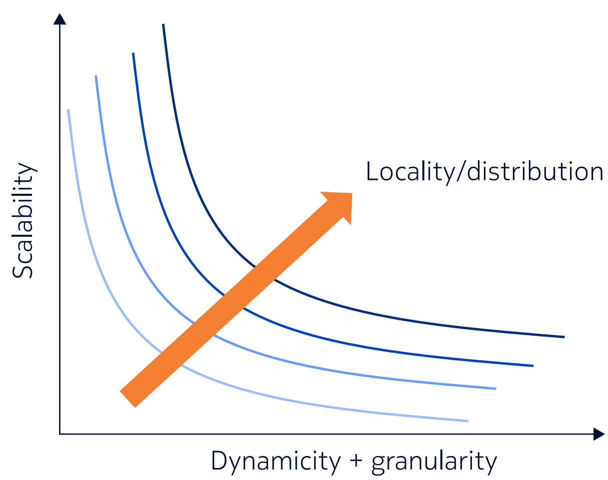
A network stripe can span across multiple domains, both administrative and technological. Its instantiation proceeds through the following three steps. First is the profiling of the network stripe, which converts the intents and contextual attributes of the application session into requirements for functional capabilities and network resources. Second is the composition of the network stripe, where the functions and resources that satisfy the requirements are identified, discovered, provisioned and interconnected. Third and last is the operation of the stripe, where all functions and resources are allocated for their intended contributions to the perceived behavior of the stripe. Progressive run-time adaptation and/or additional instances of execution of one or more of the steps may be required while the application session is running, in response to changes in the state of an application endpoint or of a network domain.

At present, the scope of the network stripe definition is limited to its outcome and does not specify how it is obtained. While implementation details may emerge over time, there is not a single way of creating a network stripe. Instead, the implementation of each instance depends on the application, the context of the session, and the network domains it involves. The profiling of the stripe is an instance of the broader intent-based management framework of UNEXT, with emphasis on application-specific requirements and performance metrics. The construction of the network stripe and then the control of its component functions follow the general UNEXT model for creation of the network-compute continuum and for autonomous service composition and operation across multiple compute and network domains, empowered by pervasive, domain-tailored artificial intelligence (AI) models.

Principles of symbiotic design

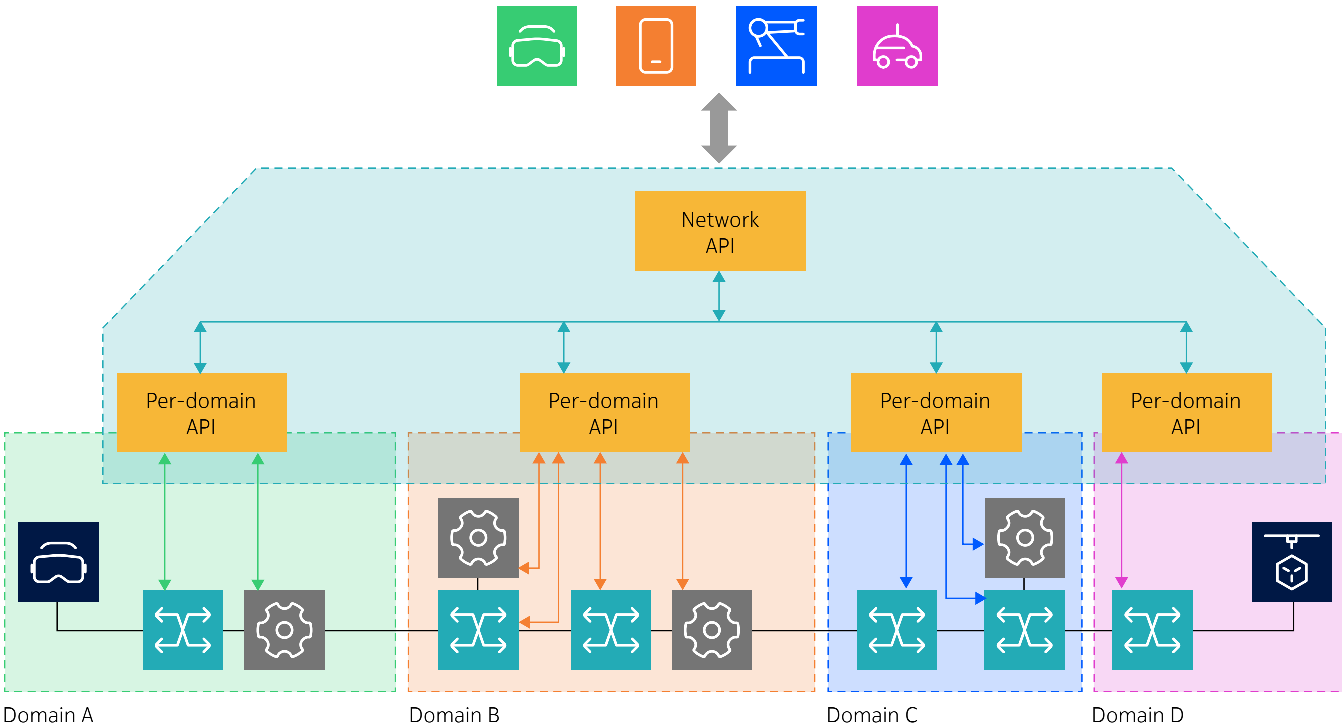
Essential to the realization of all symbiotic technologies is the observation that mutual awareness, dynamicity and granularity are inversely proportional to the size of their enforcement domain. In other words, a centralized system for making symbiotic decisions is much more opaque, slow and coarse, and therefore much less efficient than a distributed one. A network stripe that supports an application session over multiple network domains achieves its best behavior when each domain autonomously handles its interaction with the application through its own cycle of profile-compose-operation steps. Figure 3 provides a qualitative representation of the inverse proportionality that exists between scalability (and feasibility) on the one hand and dynamicity and granularity on the other. It also underscores the facilitating role of locality and distribution in the realization of highly dynamic and granular symbiotic functions, i.e., distributed enforcement results in better scalability as well as finer dynamicity and granularity.

Figure 3. Scalability is "inversely proportional" to dynamicity and granularity, but locality and distribution can ease the realization of highly dynamic and granular symbiotic functions.



The distribution of the methods for amplification of the symbiotic dimensions of awareness, dynamicity and granularity must be reconciled with the goal of complete abstraction of the network capabilities when they are offered to the application developer. The developer should be allowed to expect that the capabilities are always available and not worry about where and how they can be found. For developers that rely on that expectation, the network should present to the application a single point of communication derived from a universal intent-based application programming interface (API) that is recognized by every domain of the network stripe. Actual support for the capabilities presented by the API may vary across network domains based on their individual features, but the means of intent-based communication are always the same. On top of all other benefits, this ensures full portability of the applications from one UNEXT domain to another and over sets of UNEXT domains. At the same time, developers with a deeper understanding of the technology domains that compose the network paths of their applications should still be allowed to leverage their knowledge through distributed APIs. Both centralized and distributed API options illustrated further below are feasible in UNEXT.

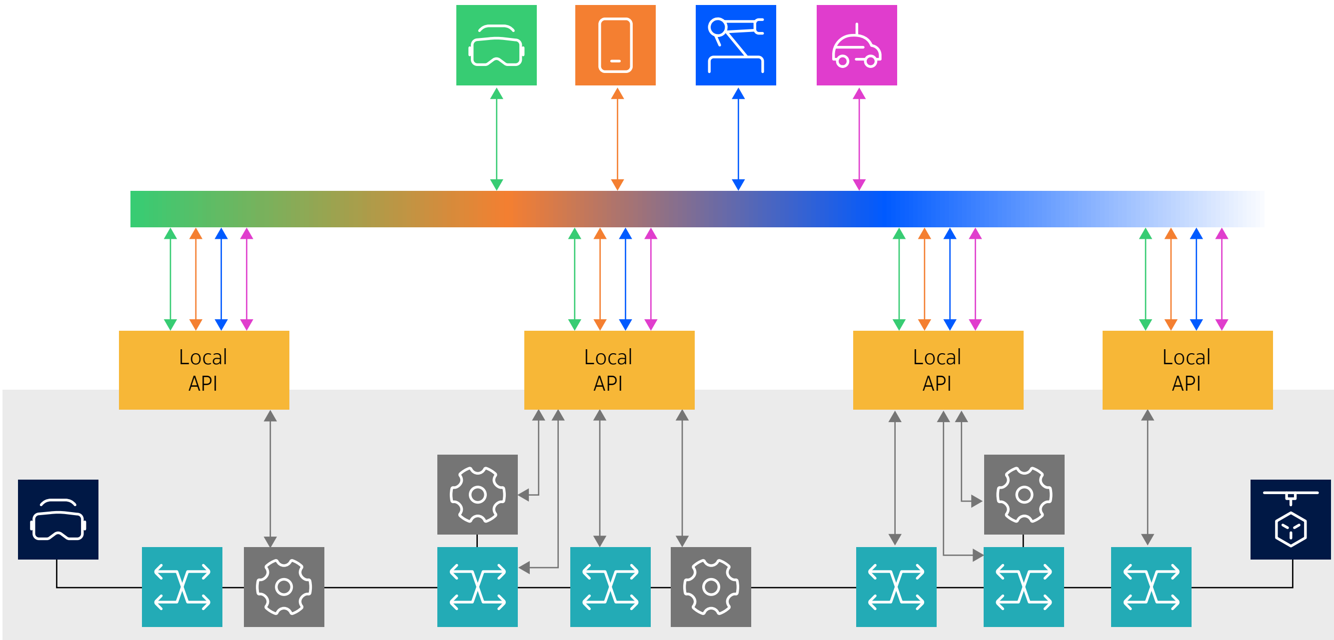
Figure 4. Hierarchy of (intent-based) APIs for applications that consume their stripes through a single API access point (the Network API).



UNEXT may resort to a hierarchy of APIs, as shown in Figure 4, to present the distributed, multi-domain realization of the stripe capabilities as a monolithic entity to applications that expect to consume the stripe through a single API access point. Moving from top to bottom in the hierarchy, every layer brings the intent formulation closer to the properties and context of the compute and network functions that compose the stripe within the respective domains.

To preserve the flexibility of allowing application developers to finely control the functions of the individual domains, UNEXT exposes all APIs in the hierarchy not only internally to the hierarchical layer immediately above but also externally to the applications that are willing to use them. Figure 5 shows this arrangement, where applications utilize local, distributed APIs for interaction with portions of the UNEXT system. Compared to Figure 4, domain boundaries have been removed in Figure 5 because the distributed APIs may be exposed not only per domain but also per function or per group of functions.

Figure 5. Network-application symbiosis through distributed APIs.



Having established the means for direct communication between UNEXT and the applications, through either a single, centralized API access point or multiple, distributed APIs per domain, the network-application symbiosis enriches the set of APIs by developing new network-compute functions and capabilities and the interfaces that expose and control them. Every new symbiotic function and capability should provide the right level of granularity for the needs of the applications and minimize the latency of adaptation to any contextual change that it recognizes for every session it supports.

Creation of network-application symbiosis

To create an efficient symbiotic environment between the application and the network, UNEXT provides functional capabilities for distributed realization and unified abstraction that deliver fine-grained support for contextual application requirements with minimum adaptation latency. These capabilities can logically be grouped by affinity with the fundamental steps in the lifecycle of the network stripe, as described in the following subsections.

Symbiotic awareness

Mutual contextual awareness with respect to requirements, capabilities, status and performance is a necessary condition for the reciprocal adaptation between the network and the application session. In this space, UNEXT principles include the definition of interfaces for the explicit exchange of information and control commands in both directions, with transparency that increases with the established level of trust between the two sides. When the interface is completely opaque, either intentionally or because means for direct communication are not yet available, UNEXT can resort to machine-learning tools to infer properties and needs of the application session and trigger adequate orchestration and control decisions. Some of the key enablers of symbiotic awareness are the intent-based expression of application requirements and performance, the dynamic learning and creation of network-application interfaces, the use of application-specific protocol headers in data packets [2], the optimization of RAN performance through quality of experience (QoE)-aware network-application interactions [3], and the run-time profiling and identification of over-the-top applications [4, 5, 6].

To further explain how symbiotic awareness helps both applications and networks, consider the run-time profiling and identification of over-the-top applications. It can be demonstrated that using sophisticated machine learning techniques, it is possible to profile and identify an application just from the system and network footprint left by its execution. When an application is identified, UNEXT can configure RAN slices in a wireless network or compute and networking resources in a data center, leveraging the interplay between the application's Quality of Service (QoS) and QoE parameters.

Symbiotic orchestration

Using all the contextual information gathered from the application session, which also keeps evolving as the session progresses, UNEXT must rapidly identify, locate, provision and interconnect all the resources and capabilities that together meet the resulting requirements. Among the general principles and methods for the orchestration of UNEXT resources, network-application symbiosis defines those that assure the speed and granularity of execution of the orchestration tasks and enable the detection of meaningful contextual changes (during the stripe setup and deployment phase). Important capabilities required for the symbiotic orchestration and developed within UNEXT include the orchestration of resources by sharing a single source of trust for user authentication across applications [7], the adaptive provisioning of RAN resources based on application requirements and performance [8], and the dynamic allocation of ultra-low-latency network paths for distributed applications [9].

To further elaborate how symbiotic orchestration of ultra-low-latency network paths for distributed applications can be realized, consider a centralized controller that receives flow specifications, assigns explicit routes to the deterministic flows, and configures traffic shapers for the ingress links based on the guaranteed rates and routes of the deterministic flows that they handle. The symbiotic ingress shapers of UNEXT constrain their scheduling decisions with routing information to minimize their contribution to the accumulation of queuing latency through the network.

Symbiotic control

Once UNEXT resources and capabilities are identified and interconnected for support of an application session, their operation must be continuously tuned in response to new contextual signals received from the session. The goal is to maintain the quality indicators of the session within the desired range as given by the application's intents while minimizing the consumption of network and compute resources allocated in the operation of the network stripe (i.e., during the operational phase of assuring the intent). Various application-specific functionalities are needed to achieve the symbiotic control capability. As practical use cases, low-latency mobile applications can be enhanced by on-demand routing optimization and mobility management [10, 11], augmented/virtual reality applications can enjoy better real-timeliness through tailored congestion control [12, 13], and one-to-many data dissemination applications like publish-subscribe can use network assistance for handling data streaming/protocol diversity and user dynamism [14, 15].

The principles of symbiotic control can be elaborated using a publish-subscribe application that leverages network assistance. In UNEXT, the control and data planes of a legacy publish-subscribe system are separated, and the protocol is extended to allow a user client to connect and express its requirements, both at setup time and during run time. Based on the client's and stream's requirements, the client can be connected (and reconnected) to the most appropriate data dissemination node within the network, whether existing or created on-demand.

Conclusions

This paper has examined network-application symbiosis within the UNEXT framework, showcasing how networks and applications can mutually benefit through fine-grained exchanges of state and requirement information. By integrating reciprocal awareness, dynamic adaptability, and granular control, UNEXT shifts from isolated operation to a cooperative model that meets the needs of modern applications.

Using AI/ML-driven services, intent-based APIs, and decentralized orchestration, UNEXT dynamically manages network resources to optimize application performance and user satisfaction. The concepts of network stripes and symbiotic orchestration ensure seamless integration across network domains, supporting next-generation applications like autonomous driving and virtual environments with the required reliability, low latency, and security. UNEXT presents a path to a more unified, adaptive, and intelligent network ecosystem, where symbiosis between applications and networks becomes essential.

References

1. M. Gruber, A. Sefidcon, C. Vulkán, "UNEXT: A Unified Networking Experience," Nokia Bell Labs White Paper, December 2023. https://www.bell-labs.com/institute/white-papers/unext-a-unified-networking-experience/.
2. L. Hao, B. Erman, A. Francini, B. Cilli, E. Bastug, C. Di Martino, "uNPE: Unified Network Protocol Encapsulation for Highly Transparent Future Networks," Proc. of IEEE ICC 2023 Workshops, Rome, Italy, May 2023. https://doi.org/10.1109/ICCWorkshops57953.2023.10283584.
3. A. Kak, V.-Q. Pham, H.-T. Thieu, N. Choi, "IntegRAN: Towards Vertical and Horizontal Integration for Programmable RAN Performance Optimization," Proc. of IEEE INFOCOM 2024 Workshops, Vancouver, BC, Canada, May 2024. https://doi.org/10.1109/INFOCOMWKSHPS61880.2024.10620794.
4. A. Kak, H.-T. Thieu, V.-Q. Pham, R. K. Sheshadri, N. Choi, Y. Guan, M. Yin, T. Han, "AweRAN: Making a Case for Application-Aware Radio Access Network Slicing," Proc. of ACM WiNTECH 2023, New York, NY, USA, October 2023. https://doi.org/10.1145/3615453.3616511.
5. C. Gutterman, E. Grinshpun, S. Sharma, G. Zussman, "RAN Resource Usage Prediction for a 5G Slice Broker," Proc. of ACM Mobihoc 2019, Catania, Italy, July 2019. https://dl.acm.org/doi/10.1145/3323679.3326521.
6. H. Chang, M. Kodialam, T. V. Lakshman, S. Mukherjee, J. Van der Merwe, Z. Zaheer, "MAGNet: Machine Learning Guided Application-Aware Networking for Data Centers," IEEE Transactions on Cloud Computing, Vol. 11, No. 1, pp. 291-307, Jan.-March 2023. https://doi.org/10.1109/TCC.2021.3087447.
7. S. P. Rao, A. Bakas, "Authenticating Mobile Users to Public Internet Commodity Services Using SIM Technology," Proc. of ACM WiSec 2023, Guilford, UK, May 2023. https://dl.acm.org/doi/abs/10.1145/3558482.3590181.
8. C. Liu, G. Aravinthan, A. Kak, N. Choi, "TinyRIC: Supercharging O-RAN Base Stations with Real-time Control," Proc. of ACM MobiCom 2023, Madrid, Spain, October 2023. https://dl.acm.org/doi/10.1145/3570361.3615743.
9. A. Francini, R. Miller, B. Cilli, C. Payette, C. Di Martino, "Lightweight Determinism in Large Scale Networks," IEEE Communications Magazine (Early Access), May 2024. https://doi.org/10.1109/MCOM.003.2300555.
10. M. Kodialam, T. V. Lakshman, "LASER: Learning Enhanced Segment Routing Using Link Telemetry," Proc. of IEEE HPSR 2024, Pisa, Italy, July 2024. https://doi.org/10.1109/HPSR62440.2024.10635940.
11. R. Bhatia, F. Hao, S. Benno, T. V. Lakshman, Y. Mi, "MobiSwift: Tree Embedding and Segment Routing Based Multi-Access Connectivity for Low Latency Applications," Submitted to IEEE ICDCS 2025.
12. B. Briscoe, K. De Schepper, M. Bagnulo, G. White, "Low Latency, Low Loss, and Scalable Throughput (L4S) Internet Service: Architecture," IETF RFC 9330, January 2023. https://www.ietf.org/rfc/rfc9330.pdf.
13. D. De Vleeschauwer, T. J. Wassing, K. De Schepper, C.-Y. Chang, C. Papagianni, "Application-Driven Traffic Differentiation," Proc. of EuCNC/6G Summit 2024, Antwerp, Belgium, June 2024. https://doi.org/10.1109/EuCNC/6GSummit60053.2024.10597104.

14. H. Chang, F. Hao, M. Kodialam, T. V. Lakshman, S. Mukherjee, M. Varvello, “Towards Network-Assisted Publish–Subscribe Over Wide-Area Networks,” Computer Networks, Vol. 231, July 2023. https://doi.org/10.1016/j.comnet.2023.109702.
15. V. Ziegler, H. Viswanathan, H. Flinck, M. Hoffmann, V. Räisänen, K. Hätönen, “6G Architecture to Connect the Worlds,” IEEE Access, Vol. 8, pp. 173508-173520, September 2020. https://doi.org/10.1109/ACCESS.2020.3025032.

About Nokia

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

As a B2B technology innovation leader, we are pioneering networks that sense, think and act by leveraging our work across mobile, fixed and cloud networks. In addition, we create value with intellectual property and long-term research, led by the award-winning Nokia Bell Labs.

With truly open architectures that seamlessly integrate into any ecosystem, our high-performance networks create new opportunities for monetization and scale. Service providers, enterprises and partners worldwide trust Nokia to deliver secure, reliable and sustainable networks today – and work with us to create the digital services and applications of the future.

Nokia is a registered trademark of Nokia Corporation. Other product and company names mentioned herein may be trademarks or trade names of their respective owners.

© 2024 Nokia

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

Document code: CID214399 (December)