6G-INTENSE: Intent-driven Native AI architecture supporting Network-Compute Abstraction and Sensing at the Deep Edge

Akram Boutouchent, Abdelkader Mekrache, Adlen Ksentini, Gereziher Adhane, João Fonseca, Joseph McNamara, Kostas Ramantas, Marco Palena, Marius Iordache, Renato Lo Cigno, Salwa Mostafa, Swastika Roy, Christos Verikoukis.

***Abstract*—The 6G ecosystem is anticipated to encompass a diverse array of business actors, including mobile network operators, infrastructure providers, and service providers. Existing network management systems, which traditionally map technological and administrative domains on a one-to-one basis, are inadequate for managing such a complex environment. This necessitates innovative distributed architectures with varying levels of abstraction. To address these challenges, we present 6G-INTENSE, a novel network management architecture designed to handle the distributed and heterogeneous nature of 6G through two key abstractions. First, the network-compute fabric component facilitates multi-domain mapping at the resource level by abstracting resource pools, including deep-edge nodes equipped with sensing capabilities. Second, the domain manager and orchestrator serves as a central entity for service orchestration, providing a global perspective on the infrastructure. The proposed architecture leverages the intent-based networking paradigm at all architectural layers, emphasizing its critical role in enabling autonomous networks. Furthermore, it integrates an innovative native artificial intelligence toolkit, which is pervasive across all layers.**

***Index Terms*—6G, network management systems, artificial intelligence, intent-based networking, autonomous networks.**

I. Introduction

While 5G profoundly transformed mobile networks with its evolved architecture supporting increased capacity, spectral efficiency, and flexibility, the rapid growth of data-hungry, human-centered applications (e.g., emerging social Virtual Reality (VR) and eXtended Reality (XR) paradigms collectively termed the Metaverse) will soon exceed its capabilities [1]. Consequently, both academia and industry are racing to shape the next-generation communication ecosystem, known as 6G, to drive the ongoing digitization of society. The future 6G smart networks will provide high-performance and energy-efficient infrastructure, enabling the development and deployment of next-generation internet and other services [1]. 6G will foster an industrial revolution and digital transformation, accelerating the creation of smart societies, improving quality of life, and facilitating autonomous systems, haptic communication, and smart healthcare. Achieving these objectives sustainably requires new approaches to the architecture, federation, and orchestration of telecommunications infrastructures. These new approaches necessitate multi-stakeholder ecosystems that promote synergies among Mobile Network Operators (MNOs) and owners of computational and networking resources. Such collaboration will share the extraordinary costs of upgrading from 5G to 6G while enabling new business models. These new architectural paradigms introduce unprecedented complexity due to the vast scale and heterogeneity of the orchestration domains involved. This complexity must be matched by equally capable automation capabilities. Therefore, 6G aims for the *holy grail* of pervasive Artificial Intelligence (AI), referred to as Native AI [2].

On the other hand, simplified management interactions between stakeholders and the 6G system are crucial in such architectures. Many efforts in research and standardization are currently being made to investigate approaches to achieve this. Intent-Based Networking (IBN), for instance, plays a pivotal role in enabling autonomous networks by specifying goals and constraints at a higher level to the Network Management System (NMS) [3]. IBN introduces the notion of *"intent,"* representing an abstract operational goal provided as input to the NMS. The latter generates the necessary low-level configurations to fulfill these intents. Although IBN is a relatively new term and technology, significant efforts have been dedicated to defining and standardizing it [4], [5]. In 6G architectures, IBN will be vital for managing increased complexity and ensuring seamless cooperation among diverse stakeholders.

In this paper, we introduce 6G-INTENSE, a novel network management architecture designed to address the distributed nature of 6G networks by adopting the IBN paradigm across all layers of the architecture. The key contributions of this paper can be summarized as follows:

- We propose a novel management architecture with a Na-

A. Boutouchent, A. Mekrache, A. Ksentini are with EURECOM, France (e-mail: akram.boutouchent@eurecom.fr, abdelkader.mekrache@eurecom.fr, adlen.ksentini@eurecom.fr).

G. Adhane, J. Fonseca, K. Ramantas, S. Roy are with IQUADRAT, Spain (e-mail: g.adhane@iquadrat.com, j.fonseca@iquadrat.com, kramantas@iquadrat.com, s.roy@iquadrat.com).

J. McNamara is with Ericsson, Ireland (e-mail: joseph.mcnamara@ericsson.com).

M. Palena is with Politecnico di Torino and CNIT, Italy (e-mail: marco.palena@cnit.it).

M. Iordache is with Orange, Romania (e-mail: marius.iordache@orange.com).

R. Lo Cigno is with University of Brescia and CNIT, Italy (e-mail: renato.locigno@unibs.it).

S. Mostafa is with University of Oulu, Finland (e-mail: salwa.mostafa@oulu.fi).

C. Verikoukis are with ISI/ATH and University of Patras, Greece (e-mail: cveri@isi.gr).

tive AI that facilitates Intent declaration, negotiation, and decision automation across Autonomous Domains (ADs). Moreover, joint communication and sensing is adopted as a key enabler, helping navigate the complexities and lack of reliability of the Deep Edge.
- The architecture integrates an abstraction layer over resource management. The Network-Compute Fabric (NCF) abstraction framework acts as a unifying entity for all orchestration domains. The abstraction of resources and their separation from service orchestration facilitates a much more scalable Domain Manager and Orchestrator (DMO).
- Finally, vertical service Life-Cycle Management (LCM) is delegated to tenants that can be a service provider, a vertical, or a virtual operator. Hence, this separates vertical service management from network service management.

The remaining sections of this paper are structured as follows. In Section II, we discuss the related works to our contribution. Section III presents the general 6G-INTENSE high-level architecture and discusses its main novelties and building blocks. Section IV examines the business model foreseen in 6G-INTENSE, mentioning the set of actors in the value chain. In Section V, we delve into the internal components of the main building blocks of the proposed architecture and discuss their roles and interactions. Finally, we conclude in Section VI by summarizing our work and outlining potential future directions.

II. Related Works

Several European projects are addressing the challenge of defining the next generation of networks, namely 6G networks. However, each project addresses it from different perspectives. For instance, 6G-BRICKS [6] addresses automation in the Radio Access Network (RAN) by integrating AI-driven zero-touch networks and intent-based automation. The main contributions include delivering a fully decentralized management plane that supports zero-touch orchestration of communication resources. In our contribution, we consider variant types of resources that are provided from different resource pools in administrative domains. In another context, HEXA-X II [7] is a flagship European project that aims to design the 6G platform, defining its use cases requirements, vision and implementation aspects of the 6G platform. In their recent paper [8], they discuss a comprehensive methodology for the design of a 6G E2E system that includes ten principles, a blueprint and a structured design process. The authors included a part related to the management and orchestration of 6G system, which leverage IBN as a management paradigm. However, they didn't provide a reference architecture for managing distributed 6G networks, as the main goal of the contribution was to identify key technology enablers for novel management and orchestration systems. Such a challenge can only be addressed by abstracting the underlying layers, facilitating control at a higher level, and creating hierarchical control layers. This is where our proposed architecture steps in, aligning with the vision of sustainable infrastructure sharing to encourage collaboration among all members of the value chain under a unified NCF.

The 6G-INTENSE European project is one of the stream B projects that address the challenge of defining the next generation architecture for 6G. Compared to other projects in the same scope, we are the only project that address the challenge of managing this novel complex environment while proposing to leverage IBN as well as a separation between service management and resource management. In terms of standardization activities, TM Forum is one of the leading organizations in defining the management plane for next-generation networks. Their proposed architecture [4] adopts IBN operations at all levels. Our proposed architecture follows the same approach by defining intent-management interfaces at the Tenant Management Platform (TMP), at the DMO, as well as on top of NCF. In our contribution, we go one step ahead and propose the TMP which delegate vertical services to the tenant, we also support the use of natural language for specifying requirements at the tenant level.

Previous research initiatives [9], [10] have contributed innovations that separately address each technological domain, proposing control-loop-based automation systems that support on-demand scaling of virtualized infrastructures, dynamic slicing and optimizing operational costs. Some of these initiatives represents research projects that propose novel contributions to shape the next generation of 6G networks. Authors in [9], [10] propose new IBN systems powered by machine learning for network management. However, isolated decision automation for technological domains makes it impossible to provide truly optimal solutions, maintain consistency across layers, or perform adaptations without generating conflicts. Therefore, there is a pressing need to consolidate automation control planes while leveraging declarative, active reconciliation approaches for the entire system.

III. 6G-INTENSE Innovations

A. Infrastructure Abstraction

The multi-stakeholder infrastructures envisioned in 6G, as per the *network of networks* concept of federated 6G RANs with dense, massive cells, cell-free Access Points, and Cloud-Edge-Continuum deployments, will introduce an unprecedented level of network management complexity due to the vast scale and heterogeneity of the orchestration domains involved [11]. 6G-INTENSE aims to abstract and federate all types of computational and communication resources within an internet-scale framework governed by an intelligent orchestration paradigm. This paradigm is capable of both satisfying service requests from tenants and orchestrating infrastructure resources offered by the underlying heterogeneous, multi-stakeholder resource pools. The 6G-INTENSE architecture is based on two fundamental abstractions, as represented in Fig. 1: (a) the DMO, which represents an operator's platform; and (b) the NCF abstraction framework, which exposes the underlying computational and networking resources of resource pools. The NCF leverages a unified data model to expose information regarding the underlying infrastructure to the DMO. This integration is achieved through various resource

Fig. 1: 6G-INTENSE High Level Architecture.

orchestrators that control specific resource pools. It is assumed that each DMO has a specific coverage area, which could span a country or metropolitan area; while resource pools correspond to the fundamental concept of a region and could represent a single administrative domain such as a data center, edge site, or even a smart building. Internet-scale operations can be achieved through operator federations facilitated by East/West APIs offered by DMOs. Although efforts in this direction are still in their infancy, telecom operators have been working on Global APIs for their Platforms that will enable a uniform and consistent exposure of their capabilities to MNO federations.

B. Tenant Management Platform

The 6G-INTENSE architecture facilitates the recursive deployment of vertical services by enabling tenants to deploy not only their services but also third-party vertical services. This is achieved through the delegation of certain management functionalities to the tenant by defining the TMP, which has limited visibility over the services it manages. Examples of the network management function possible at the TMP level include the LCM of vertical services and the translation of business Intents into service-level Intents understandable by the DMO. Additionally, the TMP can implement an application-level closed control loop that uses Key Performance Indicators (KPIs) collected at both the DMO level and from deployed applications, utilizing management utilities deployed with the service to monitor Service Level Agreement (SLA) satisfaction. It is important to note that the TMP has no authority to handle resources directly; all interactions pass through the DMO via Intents. By delegating network management functions to the tenant level, the complexity of the DMO is reduced compared to existing network management architectures, allowing for a highly scalable framework. Tenants' verticals can directly interact with the TMP to specify their business requirements in terms of service requests and SLAs. Additionally, tenants can interact with the underlying DMOs to request vertical services by consuming the exposed interfaces.

C. Intent-Driven Management

To navigate the extreme complexities of the aforementioned internet-scale ecosystem, 6G-INTENSE will offer full autonomy through actively reconciled Intent declaration and translation across all layers of the infrastructure. Intent-driven, continuously reconciling systems represent the state-of-the-art in modern automation architectures [12]. An Intent provides knowledge to the autonomous system in the form of expectations and constraints, guiding the system toward achieving business goals [4] [5]. Intent-driven approaches can bridge the gap between application requirements and automated resource provisioning. However, existing Intent-based approaches typically target a single automation domain with a one-to-one mapping of declarative Intents to predefined policies, or they consider business directives to be imperative in the system, assuming a centralized OSS/BSS layer. This approach has significant limitations regarding scalability and reliability. To address these issues, 6G-INTENSE adopts the TMForum automation architecture paradigm [4], which proposes the federation of ADs operating across different scopes and timescales, from the non-real-time business domain to the near-real-time service operations domain and finally to the real-time resources domain. The TMForum approach to federation is based on loose coupling via Intent APIs, as defined in the recently published TMF standard [4]. These ADs will be powered by AI/ML agents from the Knowledge and Intelligence component termed Native AI. Native AI refers to an architecture where AI/ML is pervasive throughout the different components and technological domains, leveraging federated analytics [2].

D. Native AI

6G-INTENSE delivers the rst concrete approach toward this vision by designing and implementing an AI-Native Toolkit (ANT), which relies on a family of data-driven functions broadly applicable in all 6G-INTENSE ADs. The ANT extends the Native AI concept by ensuring coordinated operation both vertically (spanning ADs) and horizontally (within ADs) among the different AI/ML-enabled components. An example of the integration of AI/ML models throughout the architecture is Intent handling. Cognitive Intent handlers in 6G-INTENSE are utilized across all ADs in the architecture (namely Tenant, Service, and Resource domains). Multiple cutting-edge AI techniques and the latest advances in Large Language Models (LLMs) are leveraged, recognizing the key role of Intents in automation architectures. LLMs, known for their ability to process and understand human language, are particularly useful for intent translation, enabling users to specify intents in natural language, as proposed in [10]. This allows intuitive interaction across 6G-INTENSE layers. LLMs can also facilitate natural language reporting, keeping users informed about network events, such as anomalies detected and resolved autonomously, in alignment with the natural-language based intent reporting [13]. Another signi cant application is resource negotiation, where LLMs interact with users in natural language to modify resource requests when initial intents are unfeasible.

Additionally, 6G-INTENSE will leverage Reinforcement Learning from Human Feedback (RLHF), which has proven instrumental in extending AI/ML-based models with human expertise. RLHF will be adopted as the main framework for resolving con icts in task prioritization at the DMO layer as part of the Intent renegotiation process (e.g., when “soft” requirements, priorities, and con icting stakeholder requirements are involved). The Fabric and Resource layers can also leverage RLHF to readapt decision-making policies, effectively tuning these policies online with new objectives as Intents are modi ed and renegotiated. Furthermore, 6G-INTENSE proposes the adoption of the Hierarchical Reinforcement Learning (HRL) paradigm to extend the capabilities of RLHF. Speci cally, HRL is leveraged to recursively decompose sub-tasks arising from the Intent handler and RLHF selection process into horizontal actions (e.g., at the Network-Compute orchestration domain) or vertical actions (across ADs) [14]. The propagation of Intent and policies across the ADs will typically generate an implicit or explicit hierarchical relationship among the different AI/ML-based functions across the architecture. As human-generated Intents decompose into sub-tasks, objectives, and decision-making policies, these must be consistent with the Intents, objectives, and policies established by the layers above them. In this setting, the framework provided by HRL ensures the downstream propagation of Intents from the DMO down to the Resource Layer, maintaining consistency and compliance with the human-generated Intents.

Fig. 2: 6G-INTENSE High-Level Business Interactions.

IV. Business Perspective

The distributed nature of the 6G-INTENSE approach promotes the integration of new stakeholders into the market. The separation of service and resource orchestration, in addition to the introduction of the TMP results in new actors that didn't exist in previous generations. In 6G-INTENSE, we may see stakeholders that own only DMO or TMP or only resources, which is completely different from the current 5G market, where, generally, network operators manage services using their resources. In Fig. 2, we illustrate the interaction between envisioned 6G-INTENSE stakeholders, which include verticals, TMP, DMO, NCF, and resource providers. As stated earlier, the business model proposed by 6G-INTENSE introduces new opportunities to unlock the 6G market by enabling the entry of new actors, such as service owners, DMO owners, and NCF owners. In this model, a TMP and DMO can be a single entity or separate entities, while the NCF owner may also own DMO or 6G resources. In what follows, we will depict each actor and discuss its role and interactions with other value chain actors.

An infrastructure provider or resource provider is the entity that owns 6G resources. It can possess one or various types of, these resources. In 6G-INTENSE, we envision diverse resource types: (a) Computing, which includes central cloud, edge cloud, extreme edge, and Deep Edge; (b) Networking, featuring compute interconnection networks such as SD-WAN; and (c) Radio, encompassing 6G Radio, Wi 6, and Satellite. The business interaction between infrastructure providers and the NCF is conducted on a Business-to-Business (B2B) basis. Resource usage can follow the established cloud model, offering pay-as-you-go options or long-term business agreements.

In this context, the NCF is one of the important components of the 6G-INTENSE system, as it contributes to the separation of resource and service management, which is one of the key innovations of 6G-INTENSE. The NCF abstracts the 6G resources to the DMO and is responsible for enforcing a 6G service. The NCF can be an independent entity separate from the DMO and resource providers. It can represent a virtual infrastructure provider or a resource broker that operators interact with to provision resources. It has B2B agreements with DMO(s) and resource providers, and it can serve multiple DMOs while registering and managing the usage of resources

Fig. 3: System design of the Domain Manager and Orchestrator.

from various infrastructure providers.

The DMO is the second component of 6G-INTENSE that enables the separation of service and resource management. This entity manages the life-cycle of 6G services deployed on the 6G infrastructure. Its role involves translating vertical requests and needs, either via the TMP or directly, into a service-level mesh. The DMO enforces the 6G service-level mesh using the resources exposed by the NCF. The DMO establishes B2B agreements: (i) with different NCF(s) to utilize 6G infrastructure, and (ii) to support various TMP(s). A Business-to-Consumer (B2C) agreement is established if there is direct interaction with verticals. The DMO can belong to an independent stakeholder that represents an administrative domain, similar to a network operator in 5G.

One of the innovations of 6G-INTENSE is to reduce the complexity of the service management plane by delegating some functions to the vertical service owner. Recognizing that many verticals do not know how to manage network services, the TMP entity has been introduced. The TMP is an entity that sells 6G services to verticals without owning infrastructure or being responsible for enforcing the 6G service on top of the 6G infrastructure, as the DMO does. It establishes B2B agreements with DMO(s), enabling the DMO to delegate service-level management functions. A TMP can be seen as equivalent to the virtual mobile operator actors of 4G and 5G.

Finally, verticals are entities that aim to deploy their applications or services as a 6G service on top of the 6G infrastructure. Vertical services or applications seek to leverage the capabilities of the 6G system to innovate and enhance the final user experience. Examples include autonomous driving, the entertainment industry, smart cities, and industry 5.0. In 6G-INTENSE, the vertical establishes a business agreement with the TMP or directly with the DMO. This type of agreement covers aspects such as the required SLA, cost, coverage area, and more. The business model between verticals and TMP or DMO operates on a B2C basis.

V. System Design

A. Domain Manager and Orchestrator

While the concepts of *Operator Federation* and *Resource Pool* management are envisioned as key aspects of future 6G ecosystems. A system architecture that supports orchestration at the internet scale is required to ensure operational agility and support the deployment and assurance of new services in such a complex ecosystem. To fill this gap, 6G-INTENSE proposes a ground-breaking DMO design (illustrated in Fig. 3), offering Intent APIs towards uniform and consistent capability exposure, and service requests from MNOs, vertical application providers, and B2B customers. The DMO operates an AD with a cognitive Intent handler, aiming to facilitate (a) fully autonomous service deployment and assurance in southbound resource pools; and (b) assured resource reservation and programmability of the underlying Compute InterConnection (CIC) networks via SD-WAN overlays translating Intents into network policies that will be enforced by the NCF. This would reduce the time-to-market to a minimum while increasing revenues for all involved stakeholders.

6G-INTENSE proposes the following DMO APIs, which are leveraged to exchange Intents and Intent reports with tenants, NCF, and other DMOs: (a) *Northbound Intent APIs* that receive requests from tenants (via TMP), requesting services such as vertical service instantiation in a given coverage area, with certain requirements, and constraints; (b) *Southbound Intent APIs* that send resource Intents and receive reports on Intent fulfillment from the NCF; (c) *Westbound APIs* to expose capabilities and send/receive service requests to/from other DMOs (e.g., to request/provide service outside one's coverage area). For the aforementioned APIs, the API definition is leveraged from TMForum specifications [4].

In Fig. 4, the two main 6G-INTENSE DMO workflows for autonomous service deployment and assurance are detailed. Both workflows are actuated by the cognitive Intent manager DMO workflow for service creation and assurance. The main control loop of the DMO is responsible for topology-aware placement and optimization of services to the optimal resource pool(s) given the declarative Intents. The DMO leverages AI/ML models from Native AI component to analyze the requirements and constraints received through the Intent. This Intent, in the form of ontology, contains requirements regarding the services to be provisioned as well as the KPIs to be respected. Based on this input, the DMO performs feasibility checks and conflict resolution to validate that the Intent can be fulfilled given the current infrastructure state. In case of non-feasibility, an Intent negotiation procedure is initiated with the Intent owner (tenants or verticals). This procedure begins by assessing the available resources exposed by the LMOs. Based

Fig. 4: 6G-INTENSE Internal Work ows.

on this assessment, Agentic LLMs can be employed to engage in a negotiation with the user using natural language. The LLM negotiates by proposing alternative solutions that align with the available resources, allowing users to either accept or decline the proposals based on their needs. When con icting Intents are detected, for example, when two or more Intents request incompatible resources or service parameters, the DMO evaluates the con ict by identifying the speci c resource constraints or KPI violations. The con ict resolution process involves either adjusting the resource allocation or proposing alternative con gurations to ensure that all feasible Intents are satis ed. In some cases, users may be prompted to prioritize certain Intents over others, depending on the relative importance of the services or KPIs involved.

The ful llment of service Intents received from tenants and verticals requires the design of services composed of various Network Functions, Applications as Microservices, and other Management Functions. This responsibility is handled by the Service Design component, which maps the received requirements to a service model that includes entities and connections. In this context, the 6G-INTENSE DMO adopts the generalized concept of a Service Mesh [15], leveraging cloud-native technologies for workload packaging. This encompasses cloud-native network functions, vertical microservices, and Cyber-Physical Systems (CPS) services (e.g., for sensing). The DMO considers all pathways that link events and responses internal to the Service Mesh, optimally distributing micro-service components across the distributed, multistakeholder cloud edge computing continuum. The Native AI concept proposed are instrumental for achieving the tasks undertaken by the DMO, as the Native AI is integrated into the main Intent Manager DMO loop, which propagates Intents and objectives downstream. It is worth noting that the DMO decomposes the received Intents into multiple sub-Intents to reduce the complexity of the requirements and constraints embedded in the Intent. Service Design can address each sub-Intent and realize topology-aware placement and optimization. Additionally, the Work ow Engine further decomposes the ful llment of these Intents and sub-intents into smaller tasks. Some tasks are propagated to the NCF as resource Intents, while others leverage the East-West APIs for inter-DMO coordination. The Work ow Engine is responsible for orchestrating the ful llment of the Intent and continuously monitoring the state of its task executions.

On the other side, the “DMO loop” continuously performs cost/bene t analysis, conducting topology-aware optimization while considering changes in context and behavioral constraints such as tolerable latency, utilization, and cost, as-

Fig. 5: Network-Compute Fabric Architecture Design.

suring both new and existing services. This process involves collecting Intent reports from the NCF, which include the current state of the service, including KPIs and resource usage. Additionally, the DMO leverages the resource exposure feature of the NCF to make decisions based on the exposed abstract model of the infrastructure. The Intent manager closes the control loop using Native AI components for anomaly detection and makes proactive decisions to fulfill the Intent. These decisions made by the Intent manager influence all other components and trigger a new Intent activation pipeline.

B. Network-Compute Fabric

The vision of 6G service meshes disaggregated across multi-stakeholder resource pools adds an unprecedented level of complexity due to the heterogeneity of the orchestration domains involved that must coordinate to fulfill the tenant Intent. Failures in a single resource pool, network link, or a single function of a service mesh can unexpectedly propagate, affecting services in all adjacent domains. This is a well-known failure mode of hyperscaler platforms and would be immensely pronounced in 6G, where resource pools can aggregate multiple types of sometimes transient resources (e.g., RAN, Transport, and Compute). To address this, the NCF is proposed as a key future 6G enabler, acting as a single entity providing access to all types of resources (illustrated in Fig.5). 6G-INTENSE goes a step further by incorporating generative AI techniques to simplify interaction with the NMS using natural language [10], aiming to provide a unified Intent-based resource management API for all technological domains (Cloud, Edge, Deep Edge, RAN). In this context, LLMs are intended to be used and adapted through In-Context Learning (ICL) techniques or Parameter-Efficient Fine-Tuning (PEFT) to translate natural language intents into LMO-defined structures (e.g., ETSI's NSD for the Cloud/Edge domain). Generative AI in the NCF layer improves interaction by allowing users, regardless of technical expertise, to specify high-level intents in natural language, which the LLMs can then interpret and map into structured requests recognized by the underlying infrastructure. Conversely, the NCF layer uses LLMs to translate structured, technical responses from the LMO back into natural language for natural language-based intent reporting, making it accessible to NCF consumers.

The Abstraction Framework interacts with each resource pool using interfaces exposed by the Local Management Orchestrators (LMOs). This significantly expands opportunities for future expansion by dynamically adding or removing resources (both long-term, such as cloud resources, and volatile, such as Deep Edge devices) without requiring detailed bookkeeping from the DMO, which would otherwise create bottlenecks. While the LMO handles the details of managing and orchestrating the resources, the Abstraction Framework is responsible for fully automating FCAPS (Fault, Configuration, Accounting, Performance, and Security) in the region, resolving conflicts across LMOs from different technological domains (RAN, Transport, Compute). The NCF workflow is presented in Fig. 4. To this end, the "Fabric loop" continuously reconciles Intents received from the DMO with the state of all (Network-Compute) southbound platforms via a local deployment of the AI-Native toolkit. More specifically, the NCF Abstraction Framework offers: (a) Auto-Configuration, leveraging Generative AI to create on-the-fly runtime configurations required by the vast ecosystem of LMOs (e.g., Nephio[1]

[1]https://nephio.org/

Open-RAN[2], EdgeX Foundry[3], etc.); (b) *Self-Healing,* acting as an escalation point to handle failures and capacity exhaustion by enacting adaptation primitives (e.g., service migration, service operation in ”degraded” quality of service mode, or service termination and escalation to the DMO via Intent reports generated by the Native AI HRL mechanisms); (c) *Self-Optimization,* detecting and addressing conflicts (e.g., due to conflicting service requirements for the multiple underlying platforms) and processing Intent reports where potential performance degradation (e.g., due to workload variations) or deviation from the target SLAs is reported; *(d) Abstraction,* leveraging common data models understandable by the DMO to abstract the underlying multistakeholder complex infrastructure. The NCF will continuously expose information regarding resource usage and KPIs, allowing the DMOs to make high-level decisions without considering low-level details handled by the NCF. Following this approach, Internet-scale orchestration can be achieved, and scalability issues related to one-to-one mapping of current state-of-the-art solutions can be resolved.

VI. Conclusion

The evolution towards next-generation networks, namely 6G, is impacted by the high expectations and constraints imposed on the infrastructure. It is expected to achieve unprecedented performance and enable novel applications. Realizing such ambitions will require rethinking how network and computing infrastructures are managed. 6G-INTENSE takes the first step towards this ambition by proposing a novel architecture that pushes the boundaries and limitations of the current management architecture to a new level. In this paper, we discussed the innovations brought by the proposed architecture, as well as our vision of how it can impact new businesses with the rise of novel actors and stakeholders. This is the first step towards decentralized NMSs that enable the novel 6G ecosystem.

Acknowledgment

This work is supported by the European Union’s Horizon Program under the 6G-Intense project (Grant No. 101139266).

References

[1] W. Saad, M. Bennis, and M. Chen, “A Vision of 6G Wireless Systems: Applications, Trends, Technologies, and Open Research Problems,” *IEEE Network*, vol. 34, no. 3, pp. 134–142, May 2020, conference Name: IEEE Network.

[2] W. Wu, C. Zhou, M. Li, H. Wu, H. Zhou, N. Zhang, X. S. Shen, and W. Zhuang, “Ai-native network slicing for 6g networks,” *IEEE Wireless Communications*, vol. 29, no. 1, pp. 96–103, 2022.

[3] A. Leivadeas and M. Falkner, “A survey on intent-based networking,” *IEEE Communications Surveys Tutorials*, vol. 25, no. 1, pp. 625–655, 2023.

[4] D. A. Cooperson, “Intent in autonomous networks,” TMForum, IG 1253, 08 2022, version 1.3.0.

[5] A. Clemm, L. Ciavaglia, L. Z. Granville, and J. Tantsura, “Intent-Based Networking - Concepts and Definitions,” RFC 9315, Oct. 2022.

[2]https://www.o-ran.org/

[3]https://www.edgexfoundry.org/

[6] K. Ramantas, A. N. Bikos, W. Nitzold, S. Pollin, A. Ksentini, S. Mayrargue, V. Theodorou, L. Christofi, G. Gardikis, M. A. Rahman, A. Chawla, F. Ibañez, I. Chochliouros, D. Nicholson, Mario, Montagudand, A. Shojaeifard, A. Pagkotzidis, and C. Verikoukis, “6g-bricks: Building reusable testbed infrastructures for cloud-to-device breakthrough technologies,” in *2023 IEEE Globecom Workshops (GC Wkshps)*, 2023, pp. 751–756.

[7] Uusitalo, Mikko A. and Rugeland, Patrik and Boldi, Mauro Renato and Strinati, Emilio Calvanese and Demestichas, Panagiotis and Ericson, Mårten and Fettweis, Gerhard P. and Filippou, Miltiadis C. and Gati, Azeddine and Hamon, Marie-Helene and Hoffmann, Marco and Latva-Aho, Matti and Pärssinen, Aarno and Richerzhagen, Björn and Schotten, Hans and Svensson, Tommy and Wikström, Gustav and Wymeersch, Henk and Ziegler, Volker and Zou, Yaning, “6g vision, value, use cases and technologies from european 6g flagship project hexa-x,” *IEEE Access*, vol. 9, pp. 160 004–160 020, 2021.

[8] S. Kerboeuf, P. Porambage, A. Jain, P. Rugeland, G. Wikström, M. Ericson, D. T. Bui, A. Outtagarts, H. Karvonen, P. Alemany *et al.*, “Design methodology for 6g end-to-end system: Hexa-x-ii perspective,” *IEEE Open Journal of the Communications Society*, 2024.

[9] A. Boutouchent, A. N. Meridja, Y. Kardjadja, A. M. Maia, Y. Ghamri-Doudane, M. Koudil, R. H. Glitho, and H. Elbiaze, “Amanos: An intent-driven management and orchestration system for next-generation cloud-native networks,” *IEEE Communications Magazine*, vol. 62, no. 6, pp. 42–49, 2024.

[10] A. Mekrache and A. Ksentini, “LLM-enabled intent-driven service configuration for next generation networks,” in *NetSoft 2024, 10th IEEE International Conference on Network Softwarization, 24-28 June 2024, St. Louis, USA*, IEEE, Ed., St. Louis, 2024.

[11] M. A. Uusitalo, P. Rugeland, M. R. Boldi, E. C. Strinati, P. Demestichas, M. Ericson, G. P. Fettweis, M. C. Filippou, A. Gati, M.-H. Hamon, M. Hoffmann, M. Latva-Aho, A. Pärssinen, B. Richerzhagen, H. Schotten, T. Svensson, G. Wikström, H. Wymeersch, V. Ziegler, and Y. Zou, “6g vision, value, use cases and technologies from european 6g flagship project hexa-x,” *IEEE Access*, vol. 9, pp. 160 004–160 020, 2021.

[12] A. Leivadeas and M. Falkner, “A survey on intent-based networking,” *IEEE Communications Surveys Tutorials*, vol. 25, no. 1, pp. 625–655, 2023.

[13] A. Mekrache, M. Mekki, A. Ksentini, B. Brik, and C. Verikoukis, “On combining xai and llms for trustworthy zero-touch network and service management in 6g,” *IEEE Communications Magazine*, 2024.

[14] M. A. Habib, H. Zhou, P. E. Iturria-Rivera, M. Elsayed, M. Bavand, R. Gaigalas, Y. Ozcan, and M. Erol-Kantarci, “Intent-driven intelligent control and orchestration in o-ran via hierarchical reinforcement learning,” in *2023 IEEE 20th International Conference on Mobile Ad Hoc and Smart Systems (MASS)*, 2023, pp. 55–61.

[15] W. Li, Y. Lemieux, J. Gao, Z. Zhao, and Y. Han, “Service mesh: Challenges, state of the art, and future research opportunities,” in *2019 IEEE International Conference on Service-Oriented System Engineering (SOSE)*, 2019, pp. 122–1225.

BIOGRAPHIES

Akram Boutouchent (Member, IEEE) received his engineering degree and MSc in Computer Science from the Ecole Nationale Supérieure d’Informatique in Algiers, Algeria, in September 2022. He is currently a PhD candidate in the Communication Systems Department at EURECOM in France. His research interests include Artificial Intelligence, Network Automation, and Intent-based Networking for next-generation wireless networks.