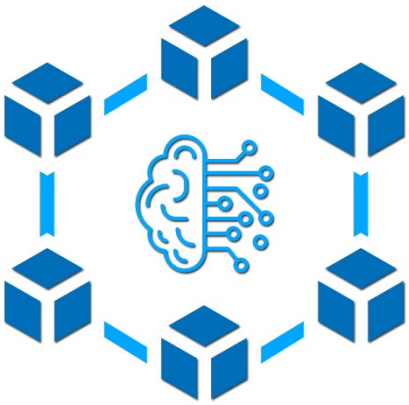

NANCY

An Artificial Intelligent Aided Unified Network for Secure Beyond 5G Long Term Evolution [GA: 101096456]

Deliverable 3.1

NANCY Architecture Design

Programme: HORIZON-JU-SNS-2022-STREAM-A-01-06

Start Date: 01 January 2023

Duration: 36 Months

Co-funded by the European Union

6GSNS

NANCY project has received funding from the Smart Networks and Services Joint Undertaking (SNS JU) under the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101096456.

Document Control Page

Deliverable Name	NANCY Architecture Design
Deliverable Number	D3.1
Work Package	WP3
Associated Task	T3.1 Overall architecture design
Dissemination Level	Public
Due Date	31 December 2023 (M12)
Completion Date	28 December 2023
Submission Date	30 December 2023
Deliverable Lead Partner	IJS
Deliverable Author(s)	Stylianos Trevlakis (INNO), Simos Symeonidis (8BELLS), Athanassios Tziouvaras (Bi2S), Rodrigo Asensio Garriga (UMU), Ioanis Makris (MINIDS), Javier Vicente (NEC), Ramón Jesús Sánchez Iborra (UMU), Gonzalo Alarcón Hellín (UMU), Noelia Pérez Palma (UMU), Juan Sebastián Camargo (i2CAT), Hatim Chergui (i2CAT), Miguel Catalan-Cid (i2CAT), Maria Belesioti (OTE), Jean-Paul Truong (TDIS), Emanuele De Santis (CRAT), Andrea Wrona (CRAT), Antonio Pietrabissa (CRAT), Alessandro Giuseppi (CRAT), Francesco Delli Priscoli (CRAT), Blaž Bertalanič (IJS), Ljupcho Milosheski (IJS), Carolina Fortuna (IJS), Jernej Hribar (IJS), Shih-Kai Chou (IJS), Daniel Casini (SSS), Alessandro Biondi (SSS), Aitor Brazaola (TECN), Marisa Escalante (TECN), Dimitrios Kavallieros (CERTH), Antonella Clavenna (ITL), Vasileios Kouvakis (INNO), Lambrini Mitsiou (INNO), Eleftherios Fountoukidis (SID), Konstantinos Kaltakis (8BELLS), Giuseppe Celozzi (TEI), Marco Tambasco (TEI), Abir Yasser Barakat (TEI), Giancarlo Sacco (TEI), Panagiotis Sarigiannidis (UOWM), Thomas Lagkas (UOWM), Dimitrios Pliatsios (UOWM), Athanasios Liatifis (UOWM), Sotirios Tegos (UOWM)
Version	1.0

Document History

Version	Date	Change History	Author(s)	Organisation
0.01	08/03/2023	Initial version	Blaz Bertalanic	JSI
0.02	28/03/2023	Adding responsibilities	Carolina Fortuna	JSI
0.03	19/04/2023	Adding to Section 3.1.6	Simos Symeonidis	8BELLS
0.04	20/07/2023	Adding to Section 3.3.2	Athanassios Tziouvaras	Bi2S
0.05	23/06/2023	Section 2	Stylianos Trevlakis	INNO
0.06	24/07/2023	Adding to Section 3.2.4 and 3.3.1	Athanassios Tziouvaras	Bi2S
0.07	24/07/2023	Adding to Section 4.3	Rodrigo Asensio	UMU
0.08	28/8/2023	Adding to Section 3.2.6	Ioanis Makris	MINDS
0.08	29/8/2023	Adding to Section 3.2.6	Eleftherios Fountoukidis	SID

0.09	08/09/2023	Adding to Section 3.3.3	Athanassios Tziouvaras	Bi2S
0.10	19/09/2023	Adding to Section 3.1.3	Javier Vicente	NEC
0.11	19/09/2023	Adding to Section 3.3.2	Ramon Jesus Sanchez Ibora	UMU
0.12	20/09/2023	Adding to Section 3.2.2	Juan Sebastián Camargo	I2CAT
0.13	20/09/2023	Adding to Section 3.2.2	Hatim Chergui	I2CAT
0.14	20/09/2023	Adding to Section 3.1.2	Ramon Jesus Sanchez Ibora	UMU
0.15	22/09/2023	Adding to Section 3.1.1	Maria Belesioti	OTE
0.16	24/09/2023	Adding to Section 3.1.1	Maria Belesioti	OTE
0.16.1	24/09/2023	Adding to Section 3.1.1	Dimitrios Kavallieros	CERTH
0.17	27/09/2023	Adding to Section 3.2.1	Ramon Jesus Sanchez Ibora	UMU
0.18.1	28/09/2023	Adding to Section 3.1.5	Jean-Paul Truong	TDIS
0.18.2	28/09/2023	Adding to section 3.2.2	Hatim Chergui	I2CAT
0.19	30/09/2023	Adding to section 3.3.1	Emanuele De Santis	CRAT
0.2	02/10/2023	Added input in Sections 3.1.4, 3.2.5 and 4	Stylianos Trevlakis, Vasileios Kouvakis, Lambrini Mitsiou	INNO
0.21	03/10/2023	Formatting and adding to Section 3.2.3. Reviewing all the content and adding comments for missing input and additional fixes.	Blaz Bertalanic, Shih-Kai Chou, Carolina Fortuna	JSI
0.22	11/10/2023	Formatting of references	Ljupcho Milosheski	JSI
0.23	03/11/2023	Added input in Section 4	Stylianos Trevlakis, Vasileios Kouvakis, Lambrini Mitsiou	INNO
0.24	03/11/2023	Added input to Section 3.1.6	Konstantinos Kaltakis	8BELLS
0.25	03/11/2023	Adding to Section 2.3	Ramon Jesus Sanchez Ibora	UMU
0.26	03/11/2023	Adding to Section 3.1.2	Ramon Jesus Sanchez Ibora	UMU
0.27	09/11/2023	Added Section 4.4.1	Alvise Rigo Anna Panagopoulou	VOS
0.28	22/11/2023	Added new content in Sections 3.2.1, 3.2.2, 3.3.2, 4.4.1	Daniel Casini Alessandro Biondi	SSS
0.29	23/11/2023	Added new content in section 3.1,5	Aitor Brazaola Marisa Escalante	TECNALIA

0.29.1	23/11/2023	Added new content in section 3.1.5 and 6	Giuseppe Celozzi, Marco Tambasco, Abir Yasser Barakat, Giancarlo Sacco	TEI
0.3	23/11/2023	Added new content in Section 1 - Introduction	Carolina Fortuna	JSI
0.3.1	24/11/2023	Added content in sec. 5	Antonella Clavenna	ITL
0.3.2	24/11/2023	Added content/figure in sec4.2	Shih-Kai Chou	JSI
0.3.3	30/11/2023	Added executive summary	Blaž Bertalanič	JSI
0.3.4	15/12/2023	Go through the comments, solve easy ones, add helper comments to help converge, trigger final adjustments.	Carolina Fortuna, Blaž Bertalanič	JSI
0.3.5	17/12/2023	Addressed comments in Sections 2, 3, 4, and 6	Stylianos Trevlakis, Vasileios Kouvakis, Lambrini Mitsiou	INNO
0.3.6	18/12/2023	Addressed the comments in Section 3.1.5	Jean-Paul Truong	TDIS
0.3.7	18/12/2023	Addressed the comments in Section 3.1.5	Giuseppe Celozzi, Marco Tambasco, Abir Yasser Barakat, Giancarlo Sacco	TEI
0.3.8	19/12/2023	Addressed the comments in Section 3.3.1	Athanassios Tziouvaras	Bi2S
0.3.9	19/12/2023	Addressed the comments in Section 3.1.2	Ramon Jesus Sanchez Ibora	UMU
0.31	20/12/2023	Addressed the comments in Section 3.1.1	Sotiris Soukaras	CERTH
0.31.1	20/12/2023	Addressed the comments in Section 3.1.6	Alekos Dimos	8BELLS
0.31.2	20/12/2023	Addressed the comments Section 3.3.1	Emanuele De Santis	CRAT
0.8	22/12/2023	QRM revisions and comments	Panagiotis Sarigiannidis, Thomas Lagkas, Dimitrios Pliatsios, Athanasios Liatifis, Sotirios Tegos	UOWM
1.0	23/12/2023	Addressed the comments from the QMR check	Blaž Bertalanič	JSI

Internal Review History

Name	Organisation	Date
Antonella Clavenna	ITL	08 December 2023
Anna Panagopoulou	VOS	11 December 2023
Alvise Rigo	VOS	14 December 2023

Quality Manager Revision

Name	Organisation	Date
Anna Triantafyllou Dimitrios Pliatsios	UOWM	22 December 2023

Table of Contents

List of Figures

List of Tables

List of Acronyms

Acronym	Explanation
AAL	Acceleration Abstraction Layers
AC	Actor-Critic
AI	Artificial Intelligence
AIOps	Artificial Intelligence Operations
AMF	Access and Mobility Management Function
AP	Access Point
API	Application Programming Interface
ASIC	Application-specific Integrated Circuits
AUSF	Authentication Server Function
B5G	Beyond Fifth Generation
BC	Blockchain
BC ADD	Blockchain Address
BFT	Byzantine Fault Tolerance
B-RAN	Blockchain Radio Access Network
BS	Base Station
COC	Cell Outage Compensation
COD	Cell Outage Detection
COM	Cell Outage Management
CoMP	Coordinated Multi-Point
CP	Control Plane
CU	Central Unit
CV	Continuous Variables
DOA	Description Of the Action
DoF	Degree of Freedom
DPR	Distributed Phase Reference
DQN	Deep Q-Network
DSA	Digitized Spectrum Assets
DU	Distributed Unit
DV	Discrete-Variable
DX.Y	Deliverable X.Y
eMBB	Enhanced Mobile Broadband
EOV	Execution Orderers Validations
ETSI	European Telecommunications Standards Institute
FCC	Federal Communications Commission
FFNN	Feed Forward Neural Network
FFT	Fast Fourier Transform
FIPS	Federal Information Processing Standards
FL	Federated Learning
FPGA	Field Programmable Gate Arrays
FSCD	Fast Smart Contract Deployment
GDPR	General Data Protection Regulation

gNB	gNodeB
HO	Handover
ISP	Internet Service Provider
KPI	Key Performance Indicator
LCN	Local Communication Nodes
LIME	Local Interpretable Model-agnostic Explanations
LOF	Local Outlier Factor
LSTM	Long-Short-Term-Memory
MA-DRL	Multi-Agent Deep Reinforcement Learning
ME	Managed Entities
MEAO	Mobile Edge Application Orchestrator
MEC	Multi-Access Edge Computing
MIMO	Multiple Input Multiple Output
ML	Machine Learning
MLOps	Machine Learning Operations
mMTC	Massive Machine Type Communication
MNO	Mobile Network Operators
MRAT-NCP	Multi Radio Access Technology Nomadic Connectivity Providers
MRAT-NCP	Multi Radio Access Technology Nomadic Connectivity Providers
NDPS	Neurally Directed Program Search
Near-RT	Near-Real-Time
NF	Network Function
NFV	Network Functions Virtualization
NFVO	Network Functions Virtualization Orchestrator
NFVO	Network Function Virtualization Orchestrator
NGN	Next-Generation Network
NG-SDN	Next-Generation Software-Defined Networking
NG-SON	Next-Generation Self-Organizing Networks
NIST	National Institute of Standards and Technology
NMS	Network Management Systems
NN	Neural Networks
Non-RT	Non-Real-Time
NSA	Non-Stand-Alone
OE	Order-Execute
OQS	Open Quantum Safe
PBFT	Practical Byzantine Fault Tolerance
PDCP	Packet Data Convergence Protocol
PoA	Proof-of-Authority
PoW	Proof of Work
PQC	Post-Quantum Cryptography
PQC	Post-Quantum Cryptography
QKD	Quantum Key Distribution
RF	Radio Frequency
RIC	RAN Intelligent Controller

RL	Reinforcement Learning
RLC	Radio Link Control
RNIS	Radio Network Information Service
RRC	Radio Resource Control
RRC	Radio Resource Control
RU	Radio Unit
SA	Standalone
SDAP	Service Data Adaptation Protocol
SDN	Software-defined networking
SDN	Software-Defined Networking
SFC	Service Function Chain
SHAP	Shapley Additive exPlanations
SIKE	Supersingular Isogeny Key Encapsulation
SISO	Single-Input Single-Output
SL	Side-link
SLA	Service Level Agreement
SMF	Session Management Function
SMO	Service Management and Orchestration
SNR	Signal-to-Noise Ratio
SOTA	State of The Art
SSI	Self-Sovereign Identity
TB	Transport Block
TD	Temporal Difference
TLS	Transport Layer Security
UDM	User Data Management
UE	User Equipment
UL	User Layer
UP	User Plane
uRLLC	Ultra-Reliable Low Latency Communication
UTXO	Unspent Transaction Output
VIM	Virtualized Infrastructure Manager
VNF	Virtualized Network Function
WDM	Wavelength Division Multiplexing
WP	Work Package
XAI	Explainable Artificial Intelligence
ZSM	Zero Touch Service Manager

Executive summary

The present document incorporates detailed descriptions of the novel NANCY Blockchain Radio Access Network (B-RAN) architecture design and its individual components. The primary objectives of the deliverable are to define the fundamental concepts and components of the NANCY architecture, its building blocks, and key innovations.

The information presented in this deliverable will lay the foundation for the tasks in Work Package 3 - "NANCY Architecture & Orchestration", Work Package 4 – "Dynamic Resource Management & Smart Pricing" and Work Package 5 – "Security, Privacy, and Trust Mechanisms". The document first provides a detailed overview of the Beyond Fifth Generation (B5G) architecture based on an existing Open-RAN (O-RAN) baseline and performs a gap analysis relative to B5G/6G objectives. Based on the gap analysis, it then provides details of the novel secure, and intelligent NANCY platform architecture, leveraging the advancements in Artificial Intelligence (AI) and blockchain technology, High-Level Network Architecture, B-RAN Architecture, and Orchestration Architecture. The report also discusses the 15 results (R1-R15) of the Description of Action (DOA), which individually address the defined gaps in energy efficiency, security, and intelligence identified in the baseline O-RAN architecture. Furthermore, it provides guidelines for each result (R1-R15) for refinement and redesign of the O-RAN components to improve and enhance the NANCY architecture. The document also provides detailed schematics of the main functionalities to support the development of the in-lab testbeds from Tasks 6.5 - "Greek in-lab testbed" and 6.7 - "Italian in-lab testbed.

To sum up, this deliverable is essential for the development of the O-RAN based NANCY architecture and its novel components. It goes into extensive detail about the gap analysis of the current O-RAN architecture and outlines the roadmap for a novel, secure, and intelligent architecture, that capitalizes on the latest advancements in Artificial Intelligence (AI) and blockchain technology.

1. Introduction

The NANCY project aims to enable personalised, multi-tenant, and perpetual protection wireless networking by introducing a secure and intelligent architecture for the (B5G) wireless network. Leveraging AI and Blockchain, NANCY enables secure and intelligent resource management, flexible networking, and orchestration. In this direction, novel architectures, namely point-to-point (P2P) connectivity for device-to-device connectivity, mesh networking, and relay-based communications, as well as protocols for medium access, mobility management, and resource allocation will be designed. These architectures and protocols will make the most by jointly optimizing the midhaul, and fronthaul. This is expected to enable truly distributed intelligence and transform the network to a low power computer. Likewise, by following a holistic optimization approach and leveraging the developments in blockchain, NANCY aims to support end-to-end (E2E) personalized, multi-tenant and perpetual protection. Finally, in order to accommodate the particularities of the new RAN that are generated due to the use of novel building blocks, such as blockchain, multi-access edge computing, and AI, a new experimentally verified network information theoretic framework will be presented.

1.1. Purpose of the Deliverable

The purpose of this deliverable is to provide the design of the NANCY architecture to achieve the aim of the project. Starting from an identified B5G architecture based on O-RAN (Section 2.2) and a gap analysis with respect to the B5G/6G objectives (Section 2.3) and in collaboration with other Work Packages (WP) (i.e., WP2, WP4, WP5, and WP6) we define the NANCY architecture and its views, particularly the NANCY Platform Architecture (Section 4.1), High-Level Network Architecture (Section 4.2), B-RAN Architecture (Section 4.3) and the Orchestration Architecture (Section 4.4). The gap analysis identified architectural gaps with respect to the baseline O-RAN architecture, gaps in terms of energy efficiency, security, and intelligence. Besides the gap identification, we also considered the fifteen results (R1-R15) proposed by NANCY as follows. First, we updated the SotA with respect to each result as general 5/6G progressed since the proposal was written. Then, we analyzed the purpose of the result and the aspects and components of the baseline O-RAN that need to be refined and redesigned from the perspective of that result, followed by a discussion related to how NANCY goes beyond the related state-of-art (SotA). Finally, we concluded the result analysis by considerations on the contribution of that result towards the realization of the NANCY architecture and presented the interconnection and dependences with other results.

Finally, we have also included 5GPPP and non-5GPPP projects in our analysis in order to identify common architectural blocks and ensure that major innovations are also considered by NANCY. The report also provides the schematics of the main functionalities at a level of detail that can be used by WP6 tasks in order to develop the in lab testbeds.

1.2. Relation to other Deliverables

This deliverable received the NANCY Requirements Analysis from D2.1 and is crucial in informing a number of technical tasks as follows. With respect to WP3, T3.2 relies also on D3.1 to deliver the common network functionalities (D3.2), T3.3 relies on it to deliver the AI-based orchestrator (D3.3), T3.4 relies on it to deliver the NANCY AI virtualiser (D3.4). With respect to WP4, T4.1 relies on it to deliver the Computational Offloading and Social-aware Caching (D4.1), T4.2 relies on it to deliver the Resource Elasticity Techniques (D4.2), T4.3 relies on it to deliver the Trustworthy grant/cell-free Cooperative Access Mechanisms (D4.3), T4.4 relies on it to deliver the Semantic & goal oriented communication schemes for beyond Shannon performance (D4.4), T4.5 relies on it to deliver the Smart Pricing Policies (D4.5). With respect to WP5, T5.1 relies on it to deliver the Quantum Key Distribution Mechanisms (D5.1), T5.2 and T5.3 rely on it to deliver the NANCY Security and Privacy Blockchain-based Mechanisms (D5.2), T5.4 relies on it to deliver the Self-healing and Self-recovery Mechanisms (D5.3), T5.5 relies on it to deliver the NANCY Explainable AI Toolbox (D5.4). With respect to WP6, T6.1 relies on it to deliver the NANCY Integrated System – Initial Version (D6.2).

1.3. Structure of the Deliverable

This deliverable is structured as follows. Section 1 provides a brief introduction, Section 2 discusses fundamental concepts of the architecture and provides a gap analysis with respect to O-RAN. Section 3 discusses the NANCY innovations, respectively how the fifteen results go beyond the SotA and realize the required functionality to fill the identified gaps. Section 4 discusses in detail the overall architecture while Sections 5 and 6 discuss the requirements for the in-lab testbeds and outdoor testbeds respectively. Finally, Section 7 concludes the report.

2. Fundamental Concepts and Components of the NANCY Architecture

Recently, blockchain has been recognized as a disruptive innovation shockwave [1], [2] Federal Communications Commission (FCC) has suggested that blockchain may be integrated into wireless communications for the next-generation network (NGN) in the Mobile World Congress 2018 [3]. In the same direction, the new concept of B-RAN was formally proposed and defined. In a nutshell, B-RAN is a decentralized and secure wireless access paradigm that leverages the principle of blockchain to multiple trustless networks into a larger shared network and benefits multiple parties from positive network effects. In more detail, B-RAN can drastically improve the network throughput via cross-network sharing and offloading [4]. Furthermore, the positive network effect can help B-RAN recruit and attract more players, including network operators, spectrum owners, infrastructure manufacturers, and service clients alike. The subsequent expansion of such a shared network platform would make the network platform more valuable, thereby generating a positive feedback loop. In time, a vast number of individual access points (APs) can be organized into B-RAN and commodified to form a sizable and ubiquitous wireless network, which can significantly improve the utility of spectra and infrastructures. In practice, the rights, responsibilities, and obligations of each participant in B-RAN can be flexibly codified as smart contracts executed by blockchain.

Considerable research effort was put into leveraging Blockchain in networks. Most published works [5], [6] and projects [7], [8] focus on the internet of things (IoT). There is also some work in cloud/edge computing [9], [10], wireless sensor networks [11], and other consensus mechanisms [9], [12]. However, only a few papers considered the future integration of blockchain in wireless communications. Weiss et al. discussed several potentials of blockchain in spectrum management [13]. Kuo et al. summarized some critical issues when applying blockchain to wireless networks and pointed out the versatility of blockchain [14]. Pascale et al. adopted smart contracts as an enabler to achieve service level agreement (SLA) for access [15]. Kotobi et al. proposed a secure blockchain verification protocol associated with virtual currency to enable spectrum sharing [16]. Finally, an early prototype to demonstrate the functionality of B-RAN was developed [17].

Bringing to fruition the notion of AI-aided blockchain wireless radio access beyond 5G networks calls for a flexible, scalable, and powerful ML-based orchestration framework, novel blockchain and attack models, a revolutionary network information theory approach, and the design of cutting-edge technology components. These include NFs for enabling common network functionalities, blockchain and cell-free radio access mechanisms, AI-based resource and network orchestration, distributed and decentralized blockchain approaches supported by MEC, and proactive self-recovery and self-healing mechanisms, as well as devising a suitable experimental-driven performance evaluation framework defined by the appropriately selected usage scenarios and relevant metrics. Additionally, NANCY will identify the critical technology gaps and invent, optimize, demonstrate, and assess the key enablers for the B5G RAN. In more detail, the NANCY approach is built upon three well-defined pillars:

- **Pillar I: Distributed and self-evolving B-RAN for dynamic scalability, high-security, and privacy in a heterogeneous environment**, by means of distributed and decentralized blockchain, PQC, as well as cell-free radio access mechanisms designs, in order to significantly improve the radio resource usage by introducing novel strategies for range/service expansion, supporting of novel use cases and killer apps, and exploiting the underutilized spectrum.

- **Pillar II: Towards the Pareto-optimal AI-based wireless RAN orchestration that maximizes energy efficiency and trustworthiness**, supports ultra-high availability and applications with diverse requirements, optimizes network topology and management, enables device collaboration as well as collaborative sensing perspective, allows system and network level AI models reproducibility and explainability, and transforms B5G RANs into intelligent platforms; thus, opening new service models to telecom/ISP and individual providers.
- **Pillar III: Distributed MEC for "almost-zero latency" and high-computational capabilities at the edge, where the data are generated,** by means of social-aware data and AI models caching and task offloading in order to transform the B5G verticals into intelligent and real-time flexible and reliable platforms.

2.1. NANCY high-level architecture

The aforementioned pillars represent the key building blocks of the NANCY architecture, which are jointly optimized and combined in order to successfully create a fully adaptive distributed network with high reliability, security, and trustworthiness. To realize this vision, NANCY utilizes the high-level architecture depicted in Figure 1. The architecture consists of three planes, namely: cloud, edge, and user.

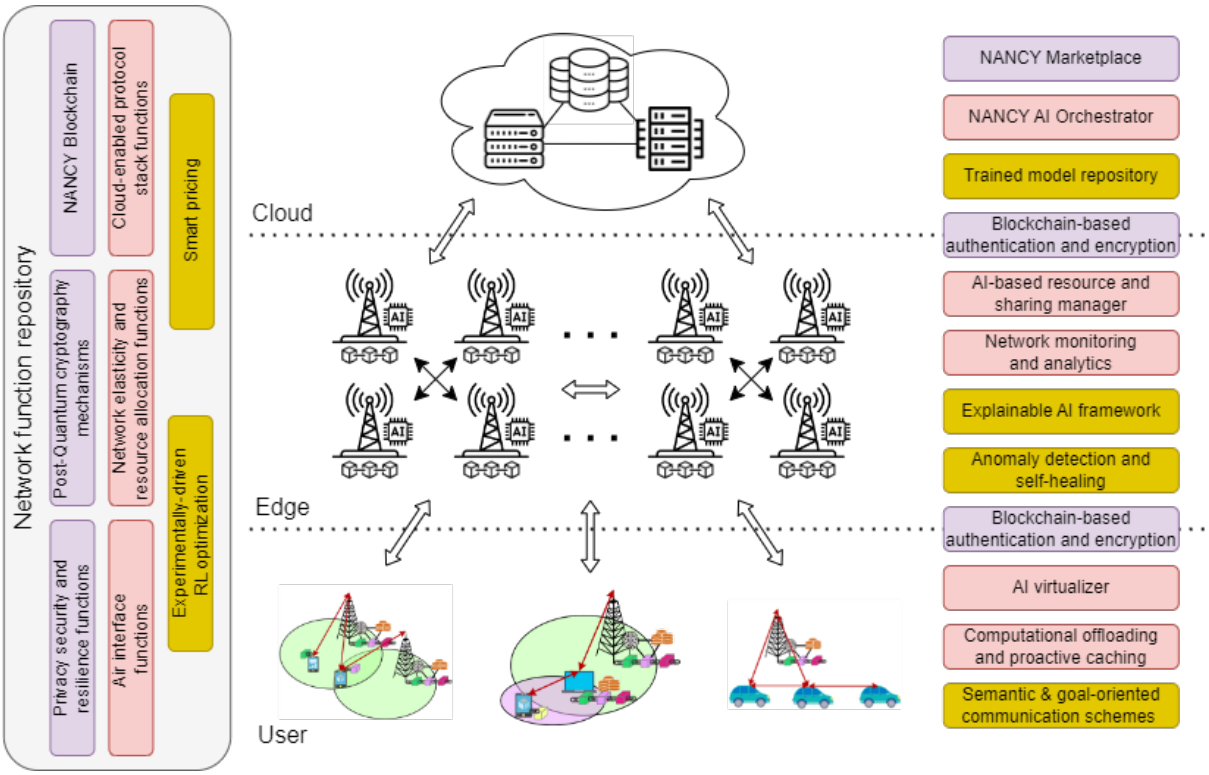

Figure 1 NANCY high-level architecture

Plane 1: Cloud plane

Several servers are equipped with strong computation, caching, and processing capabilities at the cloud plane. With a global view, this layer can leverage advanced techniques, such as data mining and big data to make a network-level orchestration shifting from reactive network operation to proactive network operation, by predicting some events or pre-allocating some resources. Thanks to the high-

computing capability and sufficient caching resources, cloud servers can process delay-tolerant applications and store content with large size or less popularity. Further, there is a central authority in the cloud plane. The central authority is equipped with tamper-resistant hardware, and it manages the security parameters and keys of all entities.

Plane 2: Edge plane

Nearby network infrastructures are geo-distributed at the network edge and equipped with MEC servers and blockchain. Network infrastructures can provide radio interfaces for mobile devices and vehicles to achieve seamless coverage and instant wireless communication. MEC servers, with computational resources, caching resources, and AI functionalities, can intelligently distribute the computational load and provide caching functionalities to support computation-intensive and delay-sensitive applications. In this plane, blockchain records all transactions generated in the wireless network and maintains a distributed ledger to increase the security and privacy of the wireless ecosystem. The transactions can be spectrum sharing between different providers, computation/caching resource allocation, energy, and storage trading, and so on.

To make a programmable, flexible, and elastic mobile edge plane, network functions virtualization and software-defined network technologies assisted by well-defined AI-mechanisms are deployed. Since network functions virtualization can abstract physical resources and establish virtual machines, the edge plane can ignore the difference in terms of vendor and protocol, and realize fast function deployment by creating, migrating, and destroying the virtual machines among distributed edge entities. Software-defined networking (SDN) can decouple network control and management functions from data forwarding such that the edge plane can apply dynamic resource management and intelligent service orchestration. Finally, this level supports anomaly detection and self-healing mechanisms.

Plane 3: User plane

In the user plane, the end-user and service provider reach an agreement on the contract terms, such as payment and resource/computational assets. These terms are explicitly recorded in a smart contract authorized by the digital signatures of the clients. The smart contract is committed to the mining network, and verified by miners to determine if the end-user has a sufficient credit balance to pay the service provider and if the service provider has enough available resources. If the contract conditions are satisfied, the verified contracts are aggregated to create a new block, which is then added to the existing blockchain. After several verifying blocks are built on top of it, the new pending block will be accepted into the main chain. The end-user will be granted time-limited access to the spectrum assets, and the service provider will automatically receive the payment for the access from the end-user. The blockchain in B-RAN can organize a large cooperative network and protect participants' benefits. The end-user's interests and the service provider's rights are enforced by the smart contracts, thereby establishing trust between an initially trustless service provider and the end-user.

NANCY Blockchain toolbox

To enhance flexibility, adaptability and programmability, NANCY architecture is further complemented with the blockchain toolbox. This toolbox is a combination of cloud-enabled protocol stack, privacy, security, and resilience as well as network elasticity and resource allocation functions. An AI-based mechanism allows the optimization of the B-RAN operation by choosing the optimal location to execute each blockchain-based functionality and selecting from the blockchain toolbox the most appropriate implementation for each case. The AI-based B-RAN orchestrator enables the flexible configuration and controlling of the B-RAN's NFs through programmable interfaces and distributed controllers as well as the establishment of the appropriate network topology for each use case.

2.2. O-RAN

The effective management and optimization of contemporary network systems necessitate the implementation of solutions that facilitate the opening of the RAN, which facilitates the disclosure of data and analytics, thereby enabling data-centric optimization, closed-loop control, and automation [18]. Presently, the methodologies employed in cellular networking are not entirely transparent and RAN components are characterized as monolithic entities, comprehensive solutions that execute all layers of the cellular protocol stack. These components are supplied by a restricted group of vendors and are perceived by operators as opaque systems. The utilization of black-box solutions has led to certain drawbacks in the RAN such as restricted reconfigurability of equipment, which cannot be fine-tuned to cater to diverse deployments and varying traffic profiles. Additionally, there is limited coordination among network nodes, which hinders joint optimization and control of RAN components [19]. Furthermore, operators are faced with vendor lock-in, which limits their ability to deploy and interface RAN equipment from multiple vendors. In light of these conditions, the task of achieving optimized radio resource management and effective spectrum utilization through real-time adaptation presents significant challenges.

In recent years, various research and standardization efforts have advocated for Open RAN as a prospective solution to address the aforementioned constraints and establish a new framework for the future of RAN. O-RAN implementations are founded on the principles of component disaggregation, virtualization, and software-based architecture [20]. These components are linked through open and clearly defined interfaces and are designed to be interoperable across various vendors. The exploitation of disaggregation and virtualization facilitates adaptable implementations, grounded on cloud-native fundamentals, which enhances the robustness and adaptability of the RAN. The utilization of open and interoperable interfaces facilitates the process of integrating various equipment vendors, thereby promoting the inclusivity of smaller players within the RAN ecosystem. Ultimately, the integration of intelligent, data-driven closed-loop control for the RAN is made possible through the utilization of open interfaces and software-defined protocol stacks.

2.2.1. O-RAN implementation

The concept of O-RAN is founded on extensive research conducted over a period of several years on open access and programable networks. The openness and programmability principles have been the focal point of the SDN revolution in wired networks over the last decade and a half. More recently, these principles have begun to permeate the wireless domain. The xRAN Forum, which is spearheaded by operators, has put forth a proposal for a standardized fronthaul interface and has also introduced the concept of open and standardized interfaces for the integration of external controllers in the RAN [21]. Simultaneously, the Cloud RAN (C-RAN) structure, which has been advocated by the operator-led C-RAN Alliance, has surfaced as a viable option to centralize the majority of the baseband processing for the RAN in virtualized cloud data centres [22]. These data centres are connected to remote radio units via high-speed fronthaul interfaces. C-RAN technology has facilitated the implementation of advanced signal processing and load balancing methods by utilizing centralized data and control paths, while simultaneously reducing expenses through the multiplexing of computational resources. The O-RAN Alliance was established in 2018 as a result of the collaboration between the two aforementioned initiatives with a primary objective to define and ultimately establish a standardized architecture and interface framework that can facilitate the implementation of an O-RAN [4]. Within a span of four years, the O-RAN Alliance has experienced a significant expansion, with its membership and contributors surpassing the 300 marks, while it is anticipated that its specifications will be responsible for generating 50% of RAN-based revenues by 2028 [23].

NANCY will adopt the basic architecture of the O-RAN Alliance, which expands upon the 3GPP NR 7.2 split employed in base stations [24]. The aforementioned approach involves the decomposition of base station functionalities into three distinct units, namely the Central Unit (CU), the Distributed Unit (DU), and the Radio Unit (RU). Furthermore, O-RAN establishes a connection between intelligent controllers via open interfaces, facilitating the transmission of telemetry data from the RAN and enabling the deployment of control actions and policies. The O-RAN framework comprises two RAN Intelligent Controllers (RICs) that execute network management and control operations at near-real-time and non-real-time time scales [19]. The O-RAN Alliance is currently in the process of establishing a virtualization platform for the RAN, while simultaneously broadening the scope of 3GPP and eCPRI interfaces to facilitate the interconnection of RAN nodes.

Over the course of the last several years, a number of open-source cellular systems designed for 5G Standalone (SA) have emerged. The research community will have the opportunity to work together with industry professionals on real-world trials via the use of these platforms, which will also help advance efforts toward the standardization and optimization of 5G SA technology. srsRAN, OpenAirInterface, and UERANSIM are three notable 5G RAN initiatives that have garnered a lot of attention from the scientific community in recent years. A particular subset of Release 16 of the 3GPP standard has been included in both srsRAN and OpenAirInterface as a result of their respective development processes. UERANSIM, on the other hand, only uses radio protocols from Layer 3, notably the Radio Resource Control (RRC) and Non-Access Stratum (NAS) layers, which comes in contrast to the other initiatives. In more detail, it does not contain Layer 2, which encompasses the Medium Access Control (MAC), Radio Link Control (RLC), and Packet Data Convergence Protocol (PDCP) levels of the RAN protocol stack. Nor does it cover Layer 1, which refers to the physical layer. Based on the above, the most promising implementations for NANCY are srsRAN and OpenAirInterface since both of these solutions provide a thorough implementation of the whole protocol stack and are aligned with the standards and specifications established by 3GPP for 5G networks.

The numerical representation of sub-carrier bandwidth, the number of channels available, and the presence or absence of accompanying documentation are the three key areas in which srsRAN and OpenAirInterface diverge significantly from one another. Increased versatility in the distribution of sub-carrier bandwidth is made possible with the deployment of OpenAirInterface. This increased flexibility makes it possible to minimize latency and makes it easier to provide support for higher-frequency bands, especially in the context of bigger bandwidths like the mmWave spectrum. It also leverages the ability to provide support for higher-bandwidths. OpenAirInterface has the potential to handle both Single-Input Single-Output (SISO) channels as well as Multiple-Input Multiple-Output (MIMO) channels, while srsRAN can only operate with single-input single-output (SISO) antenna configurations. Despite the fact that OpenAirInterface has a better degree of maturity with respect to its variety of capabilities, srsRAN offers more comprehensive documentation and community support, which allows for a deployment procedure that is more easily simplified. NANCY will employ srsRAN as an initial step towards the investigation and advancement of 5G network research based on the rationales that were presented earlier in this paragraph.

2.2.2. Building blocks

In general, it is feasible to discern four fundamental principles pertaining to the O-RAN in both scientific literature and technical implementations. These elements encompass disaggregation, intelligent controllers, virtualization, and open interfaces [25].

Disaggregation

The process of RAN disaggregation involves the separation of base stations into distinct functional components. This approach effectively incorporates and expands upon the functional disaggregation framework initially introduced by 3GPP for the gNodeB (gNB) [26]. The gNB architecture comprises three distinct components, namely the CU, DU, and RU, which are respectively referred to as O-CU, O-DU, and O-RU in O-RAN specifications. The CU is partitioned into two distinct and interdependent modules, namely the Control Plane (CP) and the User Plane (UP). The aforementioned logical partitioning enables the deployment of distinct functionalities across diverse network locations and hardware platforms. As indicated in prior literature, it is feasible to virtualize CUs and DUs on white box servers situated at the edge, with certain physical layer functionalities being hardware-accelerated [27]. Conversely, RUs are typically realized on Application-specific Integrated Circuits (ASICs) or Field Programmable Gate Arrays (FPGAs) boards, while they are positioned in close proximity to RF antennas.

The O-RAN Alliance conducted an assessment of the various RU/DU split options put forth by the 3GPP, with a particular focus on potential alternatives for the physical layer split between the RU and the DU [24]. The 7.2x split has been chosen to strike an optimal balance between the RU's simplicity and the interface's data rates and latency demands between the RU and DU. The RU in split 7.2x executes time-domain operations, including cyclic prefix addition/removal, Fast Fourier Transform (FFT), precoding, and RF operations, resulting in a cost-effective and straightforward deployment. The remaining functionalities of the physical layer, as well as those of the MAC and radio link control (RLC) layers [28, 29, 30], are handled by the DU. This includes tasks such as modulation, scrambling, layer mapping, partially precoding, and mapping into physical resource blocks. Typically, the MAC layer produces transport blocks (TBs) for the physical layer by utilizing the data that is buffered at the RLC layer, indicating that the operations of these three layers are closely coordinated. The CU units, namely CP and UP, are responsible for implementing the upper layers of the 3GPP stack. These layers include the radio resource control (RRC) layer, which oversees the connection's life cycle, the service data adaptation protocol (SDAP) layer, which manages the Quality of Service (QoS) of the traffic flows or bearers, and the packet data convergence protocol (PDCP) layer, which handles tasks such as packet duplication, reordering, or encryption for the air interface [31, 32, 33].

Intelligent controllers

The RICs constitute the second innovation, featuring configurable elements capable of executing optimization algorithms with closed-loop control and coordinating the RAN. The O-RAN vision encompasses two logical controllers that possess an abstract and centralized perspective of the network. This is facilitated by data pipelines that stream and aggregate numerous metrics pertaining to the network infrastructure's status, such as the number of users, load, throughput, and resource utilization. Furthermore, these controllers receive supplementary contextual information from sources beyond the RAN. The two RICs may utilize ML algorithms to process the data, with the aim of determining and implementing control policies and actions on the RAN. This approach entails the implementation of data-centric, self-regulating mechanisms that can efficiently enhance various network functionalities such as network and RAN segmentation, distribution of workload, seamless transfer of connections, and scheduling protocols, among other features [19]. The O-RAN Alliance has

formulated specifications for two types of RICs, namely a non-real-time (non-RT) RIC that interfaces with the network orchestrator and functions on a time scale exceeding 1 s, and a near-real-time (near-RT) RIC that handles control loops with RAN nodes on a time scale ranging from 10 ms to 1 s.

The non-RT RIC is a component of the Service Management and Orchestration (SMO) framework which utilizes the non-RT control loop to offer direction, supplementary data, and administration of AI models for the near-RT RIC [34]. Furthermore, the non-RT RIC has the potential to impact the SMO operations since it governs all the constituents of the O-RAN framework that are linked to the SMO in an indirect manner. Consequently, the non-RT RIC can make decisions and enforce policies that have a far-reaching impact on numerous devices. The issue at hand involves scalability challenges that necessitate the implementation of effective process and software design solutions.

The near-RT RIC is deployed at the network's edge, where it interacts with the DUs and CUs in the RAN, as well as with the LTE evolved Node Bases (eNBs) that comply with the legacy O-RAN standards [35]. The near-RT RIC is commonly linked with numerous RAN nodes, thereby enabling the near-RT closed-loop control to impact the QoS of a significant number of UEs. Moreover, the near-RT RIC comprises a variety of applications that facilitate personalized logic, commonly referred to as xApps, and the essential services that enable the execution of the xApps. An xApp refers to a microservice that is designed to facilitate radio resource management via dedicated interfaces and service models. The system obtains information from the RAN, such as user, cell, or slice KPIs and it may perform necessary computations and transmit control actions in response. The near-RT RIC is equipped with several components that facilitate the operation of xApps. These include a shared data layer in the form of a database that contains information on the RAN, such as a list of connected RAN nodes and users. Additionally, the platform features a messaging infrastructure that enables communication between different components and supports the subscription of RAN elements to xApps. The platform also includes terminations for open interfaces and APIs, as well as conflict resolution mechanisms that enable the orchestration of control of the same RAN function by multiple xApps.

A limitation of the current O-RAN implementations involves loops that function in the real-time domain, specifically below 10 ms, to facilitate radio resource management at the RAN node level. Additionally, these loops operate below 1 ms for device management and optimization. Real-time control encompasses various applications such as scheduling, beam management, and feedback-free detection of physical layer parameters such as modulation, coding scheme and interference recognition [36]. The loops in question exhibit a restricted scope with respect to the devices that are being optimized and are not presently integrated into the O-RAN architecture.

Virtualization

The O-RAN architecture incorporates a third principle that entails the inclusion of supplementary components aimed at managing and optimizing network infrastructure and operations. These components cover a wide range of systems, from edge systems to virtualization platforms and can be implemented on a hybrid cloud-based computing infrastructure known as O-Cloud [25]. The O-Cloud refers to a collection of computing resources and virtualization infrastructure that are consolidated in one or more physical datacentres. It integrates tangible nodes and software elements (such as the operating system and virtual machine hypervisors), as well as management and orchestration capabilities. This platform is specifically designed to cater to the virtualization paradigm of the O-RAN [37]. Benefits of utilizing this technology include the ability to facilitate hardware sharing among multiple tenants, establish uniform hardware capabilities for O-RAN infrastructure, automate the deployment and instantiation of RAN functionalities, as well as separate hardware and software components.

The O-RAN Alliance's Working Group 6 is engaged in the development of acceleration abstraction layers (AALs), which are hardware acceleration abstractions. These AALs establish shared APIs between specialized hardware-based logical processors and the O-RAN's software-defined infrastructure. This development aims to facilitate channel coding/decoding and forward error correction processes [38, 39]. Such activities are also manifested in virtualized RAN implementations that are commercially hardware-accelerated and capable of fulfilling the demands of 3GPP NR use cases, such as ultra-reliable and low-latency communications (URLLC) [40], on commercial hardware (i.e., NVIDIA [41], NEC [42], and Intel [43]) or FPGAs [44].

It is anticipated that the virtualization of RAN components and O-RAN compute elements will yield cost savings and enhance power consumption optimization pertaining to the RAN. The utilization of virtualization enables the convenient and flexible adjustment of computational resources to meet user demands, thereby restricting energy consumption to the specific network functions that are essential [45]. The closed-loop control capabilities, as well as the virtualization in the RAN, facilitate the implementation of sophisticated and adaptable sleep cycles for the BSs and RF components. These components are typically responsible for most of the power consumption in cellular networks [46].

Open interfaces

The O-RAN Alliance has recently unveiled technical specifications that delineate open and intra-RAN interfaces linking various components of the O-RAN architecture. The latter serves as a limited facilitator for the gNB disaggregated structure, which is nonetheless augmented by the O-RAN Open Fronthaul connecting the DU and the RU. The O-RAN interfaces serve to address the conventional RAN black box approach by providing data analytics and telemetry to the RICs. This facilitates various forms of control and automation actions, ranging from RAN control to virtualization and deployment optimization.

In the absence of O-RAN, the management of radio resources and optimization of virtual/physical network functions would be limited and rigid. This would result in operators having restricted access to the equipment in their RAN or resorting to a customized approach. The process of standardizing interfaces is a crucial measure in eliminating vendor lock-in within the RAN. This involves facilitating the interaction between a near-RT RIC of one vendor and the base stations of another vendor, as well as promoting interoperability among CUs, DUs, and RUs from various manufacturers. This practice also promotes market competitiveness, innovation, expedited update and upgrade cycles, and facilitates the development and implementation of novel software-based elements within the RAN ecosystem [19].

The E2 interface serves as a connection between the near-RT RIC and the RAN nodes, and is one of the interfaces that is specific to the O-RAN architecture. The RT loops are facilitated by E2, which achieves this through the transmission of telemetry from the RAN and the provisioning of control feedback from the near-RT RIC. Recently, the O-RAN alliance has started the specification of the Y1 interface to expose RAN analytics produced at the near-RT RIC and xApps to external components such as the SMO, the non-RT RIC, the 5G Core NFs or the Radio Network Information Service (RNIS) in MEC architectures. The A1 interface establishes a connection between the near-RT RIC and the non-RT RIC, thereby facilitating the implementation of policy, guidance, and intelligent models in the near-RT RIC via a non-RT control loop. The non-RT RIC serves the function of concluding the O1 interface, which establishes a connection with all other RAN components to manage and coordinate network operations. The non-RT RIC and the SMO interface with the O-RAN O-Cloud via the O2 interface. rApps hosted in the non-RT RIC have access to non-RT RIC and SMO functionalities and interfaces through the R1 interface, which also allows to expose and share rApp services. Additionally, the O-RAN Fronthaul interface

establishes connectivity between DUs and RUs. The O-RAN Alliance has established a series of specific and clearly delineated examinations to encourage compatibility among diverse interface implementations, prioritizing the fronthaul interface and E2 in its preliminary efforts.

The O-RAN architecture can be implemented through the utilization of open interfaces, allowing for the selection of various network locations (such as cloud, edge, and cell sites) for distinct equipment components. Multiple configurations for this deployment have been detailed in recent research efforts with a common practice of placing the RICs in the cloud, the CUs and DUs at the edge, and the RU on cell sites [47]. Alternative deployment strategies that involve co-locating the RICs and RAN nodes may be viable for the purpose of facilitating private and localized 5G networks.

2.3. Gap analysis

Although the fundamental principles and primary specifications for O-RAN have been formulated, allowing for the partial implementation of the use cases, numerous unresolved issues remain pertaining to standardization, development, and research. Given the openness of O-RAN sharing contexts, unique techniques are necessary to enable safe and reliable dynamic and real-time competitive O-RAN architectures.

Architecture

The O-RAN architecture is based on fundamental components such as the disaggregated RAN nodes, which consist of CUs, DUs, and RUs. Additionally, the architecture includes near-RT and non-RT RICs that host xApps and rApps, respectively. There exist several unresolved inquiries regarding the efficient implementation of this architecture, such as the optimal allocation of networking components between the edge and cloud networks, or the appropriate proportion of RAN nodes and RIC elements. Moreover, conducting additional research can facilitate the development of extensions to the O-RAN architecture for the purpose of accommodating 6G networks. For instance, the definition of dApps has the potential to facilitate instantaneous management of RAN nodes [48]. The integration of these elements with xApps facilitates the utilization of data that is not transferable for analysis from the RAN to the RICs, such as fine-grained channel estimation information or I/Q samples [49]. An additional expansion entails the amalgamation of the centralized regulation of the O-RAN framework with cell-free cellular networks, which involves a variant of massive MIMO that employs distributed antennas and centralized processing. This integration offers assistance for synchronization across the diverse antenna endpoints [50]. Several researchers demonstrate that solutions beyond current static cellular-based deployments can provide higher performance and throughput than these systems, being, e.g., Cell-Free systems, much more robust to shadow fading correlation [51]. In addition, the integration of Multi Radio Access Technology Nomadic Connectivity Providers (MRAT-NCPs) allows the optimal allocation of networking components, including RAN nodes and RIC elements to enable coverage expansion even to non-5G users.

Energy efficiency

The utilization of virtualization and closed-loop control mechanisms offer valuable foundational elements for the allocation of dynamic network functions, ultimately leading to the maximization of energy efficiency. Additional investigation is necessary to formulate orchestration procedures at the non-RT RIC and SMO that incorporate energy efficiency as a fundamental aspect of the optimization objective. Additionally, there is a need for the development of xApps and rApps that integrate control actions or policies that encompass energy efficiency targets. Efficient resource allocation with a focus on the DU provides a highly effective solution for minimizing energy consumption, while also allowing better management of bigger networks [52]. Recently, the O-RAN alliance has started the definition of

Energy Saving Use Cases which entail the utilization of rApps and xApps to optimize Energy efficiency through intelligently managing (i) carrier and cell switch on/off, (ii) RF channel reconfiguration, (iii) advanced sleep mode selection, and (iv) O-Cloud resource saving modes [52].

Security

O-RAN's level of openness expands the potential attack surface, yet concurrently facilitates novel techniques for network security. The enhanced observability of the RAN performance and telemetry, along with the ability to implement plug-in xApps and rApps for security analysis and threat identification, presents an opportunity to investigate innovative methods for fortifying and strengthening wireless networks. The advancement of security strategies that utilize O-RAN capabilities and enhance the reliability, durability, and accessibility of its implementations is a crucial measure in establishing O-RAN methodologies as a feasible and sustainable substitute for conventional RAN deployments.

Intelligence

Despite the ongoing specification of AI/ML workflows in O-RAN[1], a number of challenges persist. Firstly, the collection of training and testing datasets that are both heterogeneous and representative of large-scale deployments. Secondly, the testing and refinement of data-driven solutions through online training without negatively impacting the production performance of RAN. Finally, the design of AI/ML algorithms that can effectively operate with real, unreliable input and can generalize to different deployment conditions.

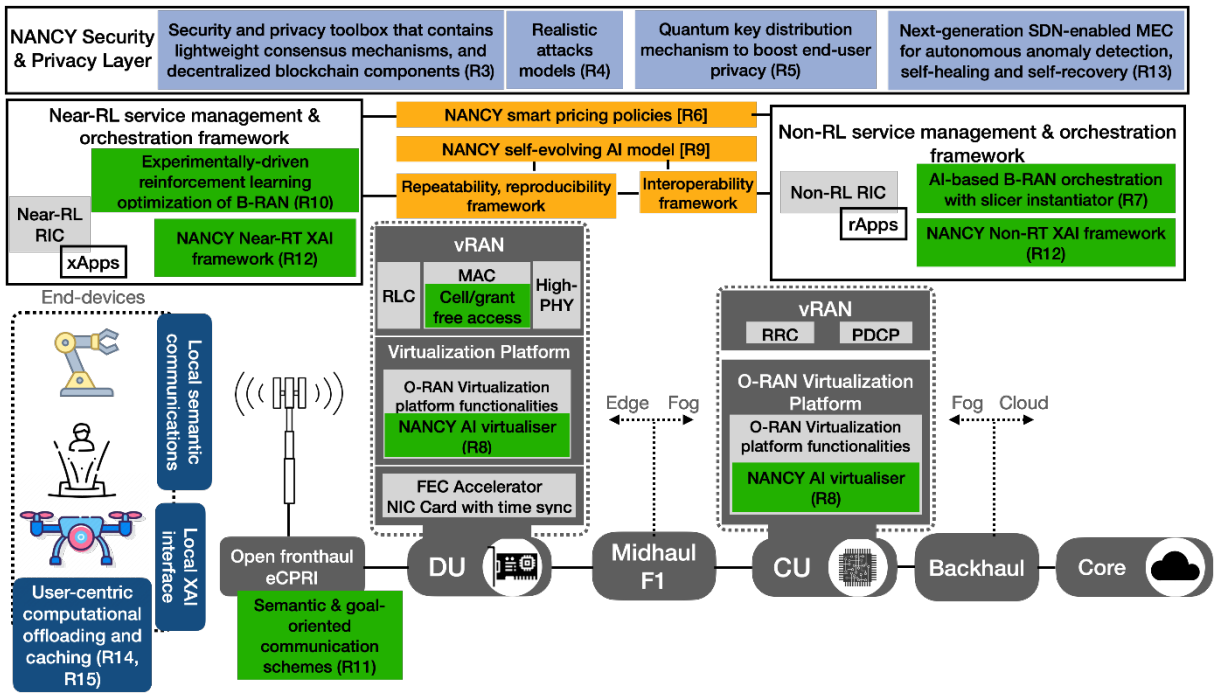

Figure 2 NANCY reference network architecture

To help promote and build confidence in these new open markets, NANCY's architecture, as depicted in Figure 2, will incorporate blockchain and smart contract technologies into O-RAN, which provide immutable and permanent records that can be audited by interested parties. The smart contract is

[1] https://docs.o-ran-sc.org/en/latest/projects.html#ai-ml-framework

used to explain the RAN user's needs and to enforce the service level agreement (SLA). In addition to the automation and network management efficiency offered by next-generation self-organizing networks (NG-SON) and O-RAN, blockchain eliminates the need for expensive intermediaries (e.g., banks, credit rating agencies) and provides unprecedented levels of transparency and trustworthiness, with the potential for significant cost savings. Furthermore, blockchain minimises the time it takes to reach agreements, allowing for true fluidity in RAN sharing. Operators dynamically sublease their resources to leverage existing infrastructure and allow other operators to enhance coverage and capacity using blockchain-enabled RAN sharing for 5G and beyond, where O-RAN is the basic architecture. Each operator is free to choose the best balance of capital investment and resource use at any time, not just when signing a RAN sharing contract or deploying a network. The democratisation and decentralisation of the telecom sector are enabled through dynamic resource trading, which allows for the creation of new competitive marketplaces with new players.

3. NANCY Enabling Innovations

In the following section, we present a comprehensive overview of the 15 results aimed to be achieved by the NANCY project, each detailed in its own sub-subsection. These results are categorically divided into three distinct yet interconnected subsections. The first subsection, "Distributed and Self-evolving B-RAN for Dynamic Scalability, High-Security, and Privacy in a Heterogeneous Environment," focuses on advancements in novel NANCY B-RAN architecture, emphasizing dynamic scalability and enhanced security. The second, "Towards the Pareto-optimal AI-based Wireless RAN Orchestration that Maximizes Energy Efficiency and Trustworthiness," highlights our strides in AI-driven RAN orchestration, balancing energy efficiency with trustworthiness. Lastly, "Distributed MEC for 'Almost-Zero Latency' and High-Computational Capabilities at the Edge, Where the Data are Generated," encapsulates the progress in edge computing, targeting minimal latency and robust computational power at the network's edge. Together, these sections paint a comprehensive picture of the NANCY project's groundbreaking aim at advancements for B5G networks.

Each sub-section includes a description of the results expected to be delivered by NANCY. The subsections are structured in a way to present the state of the i) Technical advancements and innovations state-of-the-art at the date of writing, ii) Positioning with respect to the gap analysis, iii) Technical advancements and innovations progress state-of-the-art, iv) Towards the realization of the NANCY architecture and v) Interconnection with other results.

3.1. Distributed and self-evolving B-RAN for dynamic scalability, high-security, and privacy in a heterogeneous environment

3.1.1. [R1] B-RAN architecture

Overview of the technical advancements and innovations state-of-the-art at the date of writing

In a recent study presented by Giupponi et al. [53], an integration of blockchain technology into RAN that enables Mobile operators and others to exchange RAN resources (e.g. infrastructure) in the form of VNFs autonomously and dynamically is introduced. Particularly, they define a novel O-RAN-based blockchain-enabled architecture that allows automating RAN sharing procedures through either auction or marketplace-based mechanisms. As a result of the proposed integration mobile networks will be more robust, trustworthy, wherein bringing confidence to O-RAN environments. In a complementary effort, a blockchain-enabled mutual authentication architecture for Open RAN was presented by Xu et al. [54]. In order to achieve secure and effective mutual authentication, the authors state that Blockchain Addresses (BC ADDs) issued from users' public keys are implemented. This distributed ledger approach, which offers a potential replacement for identity management in RAN without relying on third-party Certificate Authorities or Public Key Infrastructures, requires significantly fewer signaling steps and bytes for cryptographic operations than conventional techniques like Internet Key Exchange version 2 (IKEv2) and Transport Layer Security (TLS) 1.3. Using the blockchain-based RAN as a foundation, Velliangiri et al. [55] discussed a privacy-preserving framework of blockchain-enabled RAN for increased efficiency and enhanced security. The authors of this study simulated their system model in Hyperledger Fabric 1.2-based simulator, showing that the blockchain-enabled RAN achieves higher throughput and lower resource consumption compared to conventional RANs. Especially in neglected rural areas, B-RAN (Blockchain-based Radio Access Network) architecture can improve greatly the performance and sustainability of 5G networks. According to the authors of the study [56], a B-RAN architecture proposed by them operates on a consortium blockchain and has mobile network operators (MNOs) and local

communication nodes (LCNs) as network participants. The architecture's fundamental component is the innovative Proof-of-Connection (PoC) consensus algorithm, which incorporates the benefits of Proof-of-Authority (PoA) and Proof-of-Stake (PoS) processes to achieve improved effectiveness. The installation of specialized mining software on Network Management Systems (NMS) and strict node authentication are crucial components. Local digital asset staking is also necessary. The system's security and stability are guaranteed by participants' continuous monitoring of their own conduct to prevent any hostile actions.

Purpose of the component development in view of gaps defined in 2.3

B-RAN architecture is described as a proprietary way of coping with the challenges raised by the gap analysis in section 2.3. A route to increase O-RAN capabilities and handle the subtleties of 6G networks is provided by B-RAN architecture. Understanding the intricacies of B-RAN may improve O-RAN's flexibility in response to the specific requirements of 6G technology. Through tailored enhancements to the O-RAN architecture, this investigation enables 6G's network needs. Moreover, innovative techniques for network security can be made feasible by the incorporation of blockchain into RAN architecture. The architecture's objective is to improve rural areas' lack of connectivity while providing 5G networks with ultra-reliable, low-latency communication. In the aforementioned example, full nodes control ledgers, supervise smart contracts, and host decentralized apps, whilst new nodes are liable for acquiring local digital assets and registering them for monitoring. Smaller LCNs might choose lightweight node roles that only store block headers to improve the effectiveness and integrity of the system.

Technical advancements and innovations progress beyond the state-of-the-art at the time of writing

6G technology is anticipated to provide extremely high rates simultaneously with enhanced reliability and ubiquitous connectivity. In order to meet the advanced requirements and KPIs, mobile operators will need to deploy dense RANs able to serve millions of end users and mobile applications. In order to effectively address the challenges given by sophisticated requirements and KPIs as well as facilitate the growth of current networks to meet these demands, innovative approaches have subsequently emerged.

The term Blockchain refers to "a chain of interconnected information blocks forming a public ledger file for recording a list of digital actions (e.g., transactions)". These digital actions enforced by Blockchain scripts, better known as smart contracts, are powered by these actions via two steps. Firstly, smart contracts carrying digital actions are organized into blocks and broadcast to the network. Then, network nodes, also known as miners, that maintain consensus approve the transactions by inspecting the digital signature and confirming its validity. These modes create a group of valid digital actions into a new block for attachment to the end of the Blockchain via a puzzle-solving procedure known as mining.

Lately, Blockchain technology has emerged as a highly promising tool in designing scalable decentralized networks [57]. Blockchain can efficiently manage heterogeneous devices and infrastructures in 5G networks. Integrating Blockchain in RAN can be highly beneficial for operators due to its decentralized control mechanism that enables direct communications among users at the P2P level without intermediary agents, leading to lower costs and enhanced security. Blockchain mechanisms improve trust and privacy regarding large-scale, heterogeneous, and trustworthy wireless network deployment without centralized management, assisting virtual operators in integrating individually developed systems [58] [59]. Blockchain with its flexibility can be advantageous for dynamic network deployment in an operational environment having the ability to solve several issues related to trust and security in communication networks, by facilitating more efficient resource

sharing, boosting trusted data interaction, secure access control, and privacy protection, and providing tracing, certification and supervision functionalities for 5G and future 6G networks.

Beyond cutting-edge studies that implement innovative blockchain-based radio access network (B-RAN) concepts into practice are of the utmost significance in the continuously changing telecommunications sector. Improved data security, integrity, and other technological advancements that will ensure both the current and future state of telecommunications are urgently needed for the quickly evolving 5G/6G networks. The study presented by Giupponi et al. [60] proposed to adopt the Blockchain technology to enable RAN-as-a-service for future 5G/6G networks. In particular, they present a RAN as a service architecture scheme based on reverse auction mechanisms enabled by Blockchain and smart contracts technologies. Specifically, their architecture is based on two separate Blockchains supporting the aforementioned purpose. One public Blockchain (in order to achieve consensus) that records spontaneous service requests, implemented through smart contracts, for UEs that are not associated with current MNOs. Once the Blockchain registers the corresponding service delivery by operator or provider then the reverse auction mechanism is characterized as resolved. The second Blockchain is private and allows MNOs to exchange resources (spectrum, infrastructure) based on users' activity. The transaction in this Blockchain determines MNOs behavior that requests resources based on users' demand.

Additionally, according to the study [17], it is imperative to concentrate on two crucial elements in order to advance the B-RAN architecture beyond its present state. First, dynamic confirmation control can enhance network security by constantly modifying the trade-off between latency and security to address any potential trust issues inside the miner network. Furthermore, combining quantum-resistant encryption methods with trustless consensus procedures will improve security. The development of real-time Quality of Service (QoS) monitoring, blockchain-based reputation systems for Access Points (APs), smart contracts to dynamically control QoS, and decentralized auditors to publicly examine AP performance may also impact quality assurance. Together, these developments completely redefine B-RAN, promising greater security, robust service quality, and adaptability to cutting-edge technologies.

Towards the realization of the NANCY architecture

An important step towards further enhancing the NANCY architecture to successfully address the complex needs of 5G/6G networks is the integration of Blockchain technology into RAN architecture. 5G/6G networks are expected to deliver a wide range of services across several vertical sectors, with quite diverse requirements in terms of speed, latency, and capacity requiring consequently a rapid allocation of resources and network orchestration. Moreover, they are going to be highly distributed requiring the usage of technologies such as cloud-edge computing, SDN, and NFV thus increasing their complexity.

Some of the challenges posed by the aforementioned requirements can be addressed by incorporating a decentralized network like Blockchain in distributed networks like 5G/6G, delivering seamless services to end-users with transparency, security, and reliability.

While specific implementations may vary, here are some basic components and concepts typically associated with B-RAN architecture:

- RAN: The part of the mobile network responsible for connecting mobile devices to the core network infrastructure. It includes network elements such as base stations, antennas, and more.

- Blockchain: A distributed ledger technology that enables secure and transparent record-keeping through a decentralized network of computers (nodes). It ensures immutability, transparency, and consensus among network participants.
- Smart Contracts: Self-executing contracts stored on the Blockchain that automatically execute predefined actions when specific conditions are met. In the context of B-RAN, smart contracts can facilitate automated and secure interactions between network participants.
- Nodes: These are the individual computers or devices that participate in the Blockchain network. Nodes maintain a copy of the Blockchain and contribute to the validation and verification of transactions.
- Consensus Mechanism: A mechanism used by the Blockchain network to achieve agreement on the state of the Blockchain. Common consensus mechanisms include Proof of Work (PoW), Proof of Stake (PoS), and Practical Byzantine Fault Tolerance (PBFT).
- Security and Privacy: B-RAN architecture aims to enhance security and privacy in the management and operation of mobile networks. Blockchain's decentralized and tamper-resistant nature can provide protection against unauthorized access and data manipulation.

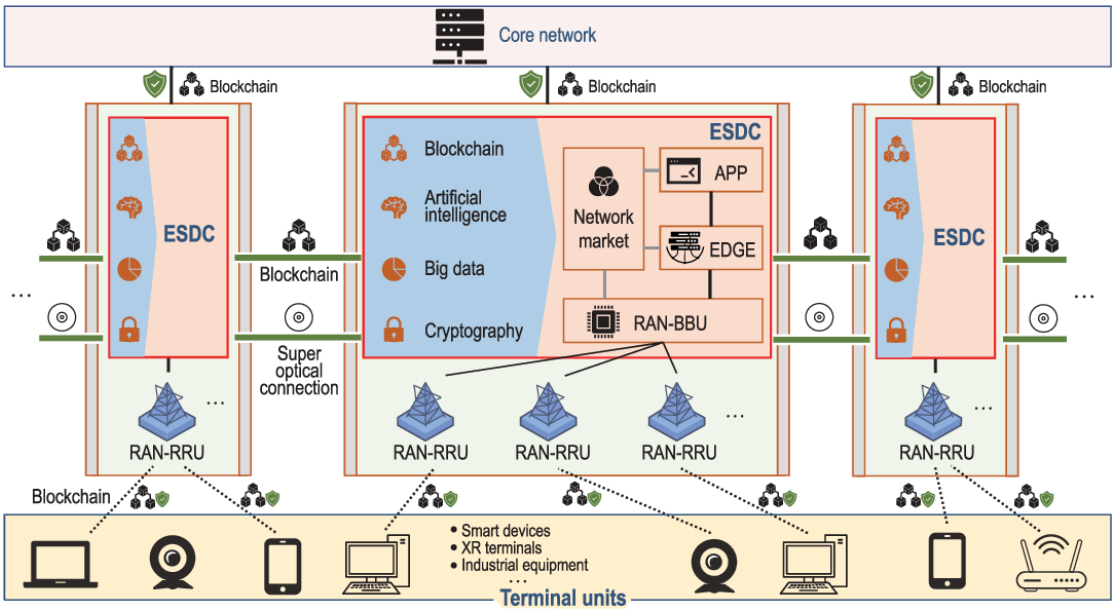

Figure 3 A promising architecture of 6G embedded with B-RAN. [61]

B-RAN can act as an open and unified framework for a variety of applications to achieve resource pooling and sharing across sectors and seems as attractive solution for emerging mobile networks. Currently, the technology has been applied in many use cases of distributed systems, such as content delivery networks [62] and smart grid systems [63]. The most essential element that guarantees data integrity in Blockchain is mining [64]. However, this process does not allow the participation of resource-limited nodes such as mobile devices, representing a significant challenge in Blockchain applications for mobile services. MEC (Mobile edge computing) architecture is used to meet the stringent low-latency requirements posed by 5G networks, becoming an auspicious solution for Blockchain RAN and applications. Moreover, it allows providers to deploy services at the edge of the network, enabling Blockchain deployment to support solving PoW puzzles, hashing, encryption algorithms, and possibly consensus.

The anticipated benefits of B-RAN include improved network reliability, reduced operational costs, enhanced privacy and security, and finally increased trust among network elements.

Interconnections and dependencies with the other results

B-RAN architecture [R1] is interconnected with the following results: [R2] Novel trustworthy grant/cell-free cooperative access mechanisms, [R7] AI-based B-RAN orchestration with slicer instantiator, [R10] Experimentally-driven reinforcement learning optimization of B-RAN and [R14] A computational offloading mechanism with novel resource-aware/provision scaling mechanisms and novel battery as well as computational-capabilities aware and offloading policies. Specifically, based on the requirements of B-RAN architecture, [R2] integrates cooperative access mechanisms as a next step. Additionally, B-RAN architecture is interconnected with [R7], since AI-based B-RAN orchestration [R7] integrates a resource orchestrator between operators. Moreover, B-RAN architecture couples also with [R10] which provides the network organization details. Finally, B-RAN is interconnected significantly to [R14] since task offloading is one of the main parts related to the architectural realization of the NANCY project.

3.1.2. [R2] Novel trustworthy grant/cell-free cooperative access mechanisms

Overview of the technical advancements and innovations state-of-the-art at the date of writing

B5G is regarded as the next stage in the evolution of communication networks, and it is expected to improve wireless capabilities to unparalleled levels. B5G technology tries to efficiently use the radio frequency spectrum and extend coverage to previously inaccessible areas or users with power limitations for transmitting at sufficient power levels [65]. The radio frequency spectrum is a limited and congested resource, making its optimization crucial. Furthermore, expanding the coverage of wireless networks poses a significant challenge. The B5G paradigm proposes developing novel spectrum access techniques and innovative solutions for coverage extension, such as low-power infrastructures, unconventional propagation technologies, and energy-efficient devices [66].

Relay stations and cell-free access technology are innovative advancements in wireless communication systems. They improve coverage, capacity, and spectral efficiency by extending signal strength and eliminating the need for traditional cellular cells [67] . These technologies improve network capacity, reliability, and coverage, supporting emerging technologies and high-bandwidth applications. The latest advances in the field offer fast data rates, controlled latency, and improved user experience [68].

Since this research topic has captured the attention of both academia and industry, significant research work has been done. In [69], authors propose an architecture for multi-hop cellular networks, which needs minimal changes in the existing network elements and interfaces. They introduce a new node called Proxy eNB (P-eNB) in the Core Network, which enables the control of multiple Relay Nodes (RNs). The RNs automatically form a multi-hop network as they are plug-and-play devices. Most of the other research works are mainly an extension of the 3GPP architectures. For example, to support mobile relays, the architecture proposed in [70] needs two LTE/EPC networks: the first network manages the UEs, while the second one is deployed by the transport operator to manage mobile relays. Nonetheless, work in [71] focuses on a heterogeneous multi-hop complementarity between cellular long-range technologies using the Uu interface and short-range side-link (SL) technologies using PC5 interface. Hence, there are two modes for SL resource selection; UEs in the first mode are within network coverage, so SL scheduling can be based on either dynamic or configured grant. However, in the second mode, UEs can operate without network coverage, which means they will autonomously handle SL radio resource management.

Purpose of the component development in view of gaps defined in 2.3

Regarding the gaps exposed in 2.3, cell-free access mechanisms aim to ensure the optimal positioning of RAN nodes, which enhances coverage and reliability of connections. This involves deploying multiple Multi Radio Access technology nomadic connectivity providers (MRAT-NCPs) throughout the service area to establish multi-hop networks when needed. Thanks to these mechanisms, UEs will be able to

move throughout the entire area while maintaining connectivity due to operator cooperation, thereby achieving greater network capacity. These multi-hop connections, enabled by strategically deployed MRAT-NCPs, ensure that data can hop from one node to another, achieving network reliability and extending coverage in challenging environments. Furthermore, the network infrastructure will provide radio interfaces for mobile devices and vehicles to achieve seamless coverage and wireless communication in areas where the fixed infrastructure does not provide acceptable connectivity.

In order to handle these dynamic interactions among end-nodes, an exhaustive control and pricing system is required. The incorporation of blockchain systems and the use of smart pricing policies facilitate interoperability between operators. Furthermore, security and trust gaps identified in 2.3 are solved through immutable records, including authentication and authorization mechanisms that integrate the use of blockchain to guarantee data integrity.

Smart pricing enables efficient resource management by establishing fees based on resource demand and availability, ensuring cost-effective allocation of network capacity to UEs from different networks. These processes are seamlessly done for the user, achieving a major quality of the user experience through collaboration between operators.

Technical advancements and innovations progress beyond the state-of-the-art at the time of writing

Through the integration of cell-free access, multi-hop networking, and Relay Nodes (RN), NANCY aims to address common challenges encountered in traditional O-RAN (Open Radio Access Network) architectures. Special focus will be given to challenges related to the optimal allocation of networking components, including RAN nodes and RIC elements. Using cell-free access mechanisms as well as nomadic connectivity, providers can help address architectural challenges, since these mechanisms enable dynamic allocation of resources and flexible deployment of RAN nodes, improving resource sharing. Additionally, the integration of these mechanisms can improve synchronization of distributed RANs, since the flexibility of MRAT-NCPs allows the agile deployment of RAN nodes strategically based on the changing needs of xApps and network conditions. Thanks to the mobility nature of MRAT-NCP, operators can increase or decrease the number of RAN nodes dynamically or allocate resources to other purposes, always adapting to the current or predicted context. Therefore, enhancing the scalability of the system.

Operators will collaborate with each other in order to enable UEs from one operator to connect to the MRAT-NCPs of another operator as required, therefore, UEs will be able to fluidly transition among different MRAT-NCPs, ensuring uninterrupted connectivity. Blockchain systems will be used to record transactions and agreements between operators, as well as to maintain secure authentication processes that guarantee data integrity and privacy. Furthermore, the integration of blockchain into the Radio Access Network (RAN) will allow the development of intelligent policies to regulate fees and resource sharing.

Contribution towards the realization of the NANCY architecture

NANCY's architecture will incorporate cell-free cooperative access mechanisms, using MRAT-NCPs distributed across the service area. The aim of this technology is to address the efficient resource allocation challenges identified in the network architecture GAP Analysis. Within the NANCY framework, multi-hop networks will be implemented to facilitate data transmission through multiple MRAT-NCPs before reaching its destination. These MRAT-NCPs serve a crucial role as signal amplifiers, improving network coverage, capacity and efficiency. This enhancement proves particularly valuable in scenarios where establishing a direct connection with the base station is unfeasible. Additionally, within multihop networks, traffic can be efficiently redirected to prevent congestion, ensuring optimal resource sharing and network performance.

NANCY will also focus on the struggle to provide sufficient coverage and latency gaps, which will be solved through the integration of multi-hop networks, allowing more efficient communication paths. Furthermore, UEs will be able to transmit data through multiple MRAT-NCPs, reducing latency issues that may impact real-time applications and user experience.

Evolve into the major challenges of cell-free mechanisms access is the ability of user equipment (UE) to have unrestricted mobility, which can be addressed through cooperation among operators. This solution enables UEs to connect to operators they do not belong to. Consequently, it can achieve uninterrupted and seamless connectivity for the user.

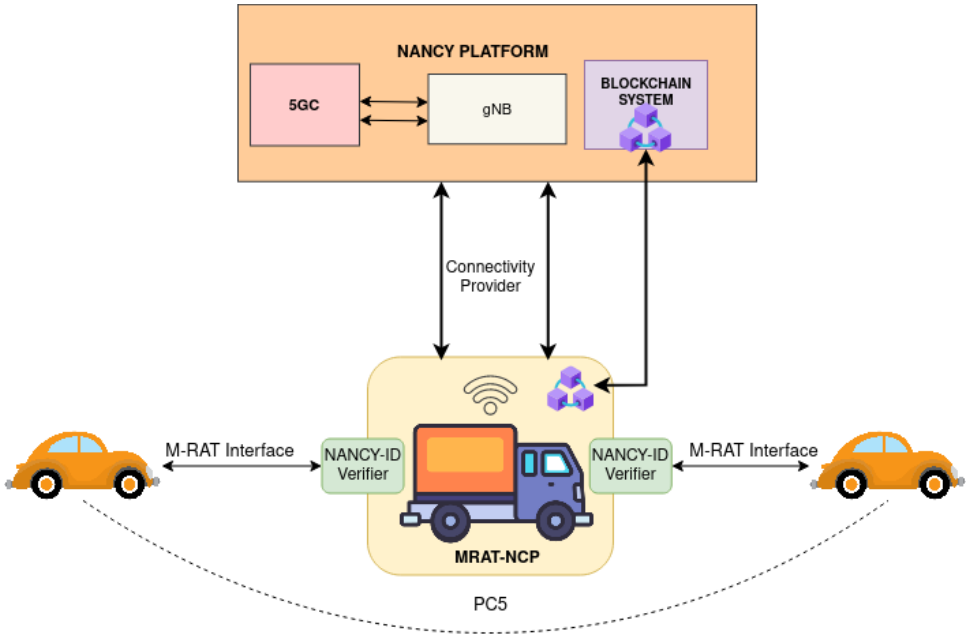

Figure 4 Connectivity through a MRAT-NCP

Interconnections and dependencies with the other results

The trustworthy grant/cell-free cooperative access mechanisms module is interconnected with the following results: [R1] B-RAN architecture, [R3] A novel security and privacy toolbox that contains lightweight consensus mechanisms, and decentralized blockchain components and [R6] Smart pricing policies. This module is interconnected to [R1] as it integrates cooperative access mechanisms such as relay nodes, taking into account the requirements of the B-RAN architecture. [R3] module provides privacy and ensures data integrity through the utilization of blockchain systems, which are used to record transactions and agreements between operators and to secure the authentication process for UEs and relay nodes. Finally, thanks to [R6] module, smart pricing policies are developed to regulate fees and optimize resource sharing within the network.

3.1.3. [R3] A novel security and privacy toolbox that contains lightweight consensus mechanisms, and decentralized blockchain components

Overview of the technical advancements and innovations state-of-the-art at the date of writing

A blockchain ecosystem contains multiple components. There is usually a core blockchain software that consists of client software that is run by a number of *peers*, who participate in the network. These peers are usually called nodes, who all have a copy of the ledger. Consensus ensures that all participants share an identical copy of that ledger. Blockchain systems also feature the so-called wallet, which represents the client or end users. Overall, a blockchain wallet is a digital wallet that allows peers to store, manage, and trade their cryptocurrencies, and which charges dynamic fees depending on cost factors (e.g. transaction sizes and others).

Just like any other software system, blockchains have different known vulnerabilities, depending on their public or federated nature, their consensus mechanisms, smart contracts platform, and other features. Very broadly, one could say that the more features the blockchain offers, the larger the attack surface becomes.

One known vulnerability, most especially in public blockchains, is an eclipse attack. Eclipse attacks can derive into many different malicious effects for the blockchain, double spending, faking the current state, temporary loss of the ledger integrity of the attacked party by means of alternative history, delaying legitimate transaction or block times, and intercepting traffic that was originally directed to a legitimate destination.

These attacks can originate from network vulnerability, but the efficiency with which they propagate and the time it takes the blockchain to recover (or not) from them, can be related to the consensus mechanism that is being used by the chain. The Proof-of-Work (PoW) mechanism, for instance, is based on solving a mathematical problem. Whichever node first solves the mathematical problem can then propose a block for validation and be rewarded with an asset (usually cryptocurrency); the more computational power, the more chances for that node to be rewarded. While it takes a considerable amount of resources to attack a PoW-based blockchain, it also can take an enormous amount of resources to retake control of the chain if attacked by a 51% attack.

From a technical perspective [72] Byzantine fault tolerance (BFT) is a generic software technique used for ordering transactions on a distributed system even if a fraction of the nodes is controlled by a malicious adversary. A BFT-based blockchain is more resilient to eclipse or alternative history attacks but at the expense of lacking scalability.

Vulnerabilities can also take place at the wallet level, where they would try to make the wallet software trust a wrong state or a wrong update by exploiting its limited knowledge and processing capability, or at the smart contract level, where they would exploit some unexpected behavior of the smart contract in some edge cases to gain advantage. Reentrancy attacks [73] leverage the state inconsistency that arises between the time funds are sent to the withdrawing party and the time the state of the smart contract is updated. By exploiting this discrepancy, the attacker is able to withdraw more than their original balance, hence stealing from the smart contract and the other users.

Purpose of the component development in view of gaps defined in 2.3

Blockchain is a rather novel technology that is constantly improving. In the case of NANCY, we propose a new combination of blockchain components able to achieve:

1. Higher security

Permissioned blockchains are typically more secure than public (permissionless) blockchains since they run based on a known consortium, and peers can be legally linked to each other. Also, due to misbehavior being not anonymous, abuse can be easily connected to one or more identifiable nodes. Overall, a permissioned blockchain restricts access to only the consortium members and can limit certain rights for certain nodes in the chain. Malicious or fraudulent nodes can be excluded from the system, which can be done through the system's governance, which is decided by the consortium members and who are responsible for e.g., issuing certificates to participate and granting/revoking rights for participation, as triggered by certain conditions.

Some implementations, like Hyperledger Fabric, enable an additional layer of control called *channels* [74] [75]. A Hyperledger Fabric channel is a private "side chain" between two or more specific consortium members, for the purpose of conducting private and confidential transactions. Each transaction on the network is executed on a channel, where each party must be authenticated and authorized to transact on that channel. Each peer that joins a channel is authenticated by its organization's certificate.

Examples of permissioned blockchains are frequent in the supply chain and logistics industries. Some primary goals in supply chain management are better transparency, traceability, and auditability of materials and products across the chain [76]. These are no different from some of the key goals in the 5G services supply chain of NANCY and are important for the stakeholders and end-users, but also for the blockchain itself since traceability and auditability are fundamental to sustaining security in the network.

Additionally, there are also off-chain mechanisms that can help against certain attacks, like front running (see NANCY D2.2): for instance, using time-stamped ordering in the NANCY marketplace and contrasting such information with on-chain counters.

2. Higher privacy

Access control as regulated by gatekeepers in permissioned blockchains helps but does not necessarily guarantee user privacy. Other mechanisms for anonymity and unlinkability should be studied. Furthermore, other perspectives on the privacy term [75] should be analyzed and improved:

- Transaction Data Privacy: Transactional activity of an entity such as a resource provider or consumer.
- State Data Privacy: Chaincode and smart contract *data* that refers to what is being offered and purchased, which peers are buying, and under which conditions.

Back to user privacy, the GDPR dimension should also be considered. Although public-private key encryption contributes to protecting confidentiality and ensuring pseudonymity-level of privacy for the users of the blockchain from the outside world, methods exist for linking individuals to public keys by analyzing blockchain transactions and other publicly available data. In particular, the GDPR [77] defines personal data broadly by setting a threshold for identification that is rather low, recognizing any means "reasonably likely to be used," considering all objective factors, such as cost, time, and

available and anticipated technology. Hence, methods that improve such anonymity and unlinkability must be studied, and the use of SSI with a PQC-able wallet is an avenue worthy of investigation.

3. Better performance

Security and privacy cannot happen at the expense of performance, but what the current out-of-the-box blockchain solutions offer might not be enough for a 5G services marketplace, for which the following features are mandatory:

a. Lower latency preserving high throughput
   Hyperledger Fabric [74] is an open-source permissioned ledger that processes transactions in three steps, relying on two types of nodes: Peers for Execution and Validations and Orderers (EOV) to establish a total order of transactions. This EOV model removes the execution bottleneck of the Order-Execute (OE) model used, for instance, in Ethereum, but it also poses difficulties and risks like failed transactions that do not add to the successful transaction throughput (goodput).
   The latency of Hyperledger Fabric is currently dominated by cryptographic verification, and several optimizations exist, like FastFabric and XOX Fabric. By removing batching and applying further optimizations, Kuhring et al. [78] achieved a reduction in latency by two orders of magnitude. However, such solutions do not include the bottleneck that a large network would require to reach a Byzantine fault tolerant (BFT) ordering.
b. Better scalability
   As previously mentioned, a BFT-based blockchain is more resilient to alternative history attacks but at the expense of lacking scalability. Scalability might not be so critical for a permissioned blockchain (i.e., a *business* blockchain, such as NANCY) when compared to latency or throughput, but still, there is room for improvement or, at least, analysis in this area. Relevant previous literature will be studied to help us design the NANCY blockchain, like Nasir et al. [79], Thakkar et al. [80] and Rüsch [81].
c. Lighter consensus mechanisms
   According to [82] from the Digiconomist's research, almost 792 kWh of electrical energy is consumed to perform just one transaction in Bitcoin (which uses PoW consensus), and that equals 376 kg CO2 of carbon footprint, which is the same as Bolivia's average electrical energy consumption per capita. Permissioned blockchains do not use PoW consensus mechanisms and hence are more efficient, especially in smaller devices like end-user terminals or IoT sensors/actuators. This improves their energy efficiency, which has stopped from being a mid- or long-term requirement, becoming an immediate necessity. Block size, block time (much dependent on the consensus mechanism), and number of transactions are important factors that must be optimized in NANCY to save processor, memory, network, and disk resources.

Technical advancements and innovations progress beyond the state-of-the-art at the time of writing

NANCY proposes several innovations belonging to result R3. Firstly, a combination of anonymous credentials (Self-sovereign Identity (SSI) will be studied) within a blockchain environment. With SSI, users can manage their identity data and decide which information they want to disclose and to whom in order to maximize their privacy level. This contrasts with single sign-on schemes or centralized login services which reach lower levels of user privacy (and security). Open-source implementations for blockchain technologies will also be studied, with a focus on Hyperledger Fabric. Fabric features several characteristics that can turn into assets for NANCY's blockchain, for instance [83]:

- Permissioned architecture
- Pluggable consensus
- Open smart contract model — flexibility to implement any desired solution model (account model, UTXO model, structured data, unstructured data, etc). Supports Turing complete smart contracts.
- Flexible approach to data privacy: data isolation using 'channels', or share private data on a need-to-know basis using private data 'collections'.
- Multi-language smart contract support: Go, Java, Javascript.
- Designed for continuous operations, including rolling upgrades and asymmetric version support.
- Governance and versioning of smart contracts
- Flexible endorsement model for achieving consensus across required organizations.

A second area of innovation under R3 is the protection of smart contracts against attacks. Two common smart contract vulnerabilities are (1) lack of verification checks in the code or code which can lead to incorrect calculations (e.g., arithmetic overflows that make the smart contract malfunction) and (2) reentrancy attacks. In a reentrancy attack, a function in the original smart contract calls another external function in a second, untrusted contract. Then the untrusted contract that is being called makes a recursive call back to the original function in an attempt to e.g., drain the funds of the initial contract or simply delay its execution until it is meaningless. Mitigation measures that can be studied include static analysis, language-based security and runtime verification.

Lightweight primitives for permissioned blockchain will be analyzed and proposed for NANCY, too.

Lastly, NANCY will provide the users with a wallet that enables them to interact with the blockchain, but most interestingly, this wallet will feature post-quantum security (please see R5).

Contributions of the result towards the realization of the NANCY architecture

The NANCY Blockchain is a fundamental element of the architecture. Before we present various statements that explain how this element interacts with others, let us explain three different roles that actors in NANCY may have within the Blockchain:

1. The *orderers* that pack transactions into blocks and agree on the order of each block.
2. The *nodes/validators/peers/endorsers* that store the full Blockchain, execute the smart contracts and serve information to the clients.
3. The *clients* (or *wallets*) that transact on the blockchain through nodes. Here, nodes are used to fetch information for the clients and execute the clients' transactions.

In NANCY, *devices* are clients (and not nodes or orderers) and the term *end-users* is referred to clients as well. Service providers (i.e. network providers) provide *nodes* to access the Blockchain, while the orderers are handled by the Blockchain consortium in the background, transparent to end-users.

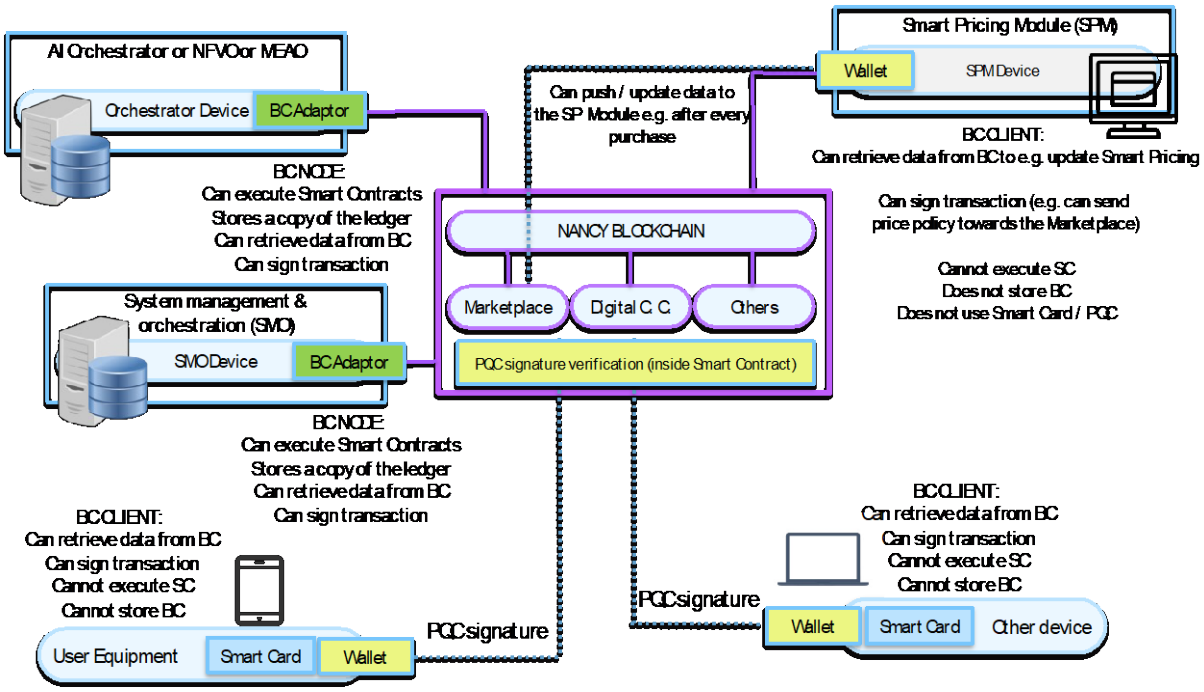

Figure 5 Key components of the NANCY architecture and their basic interactions with the Blockchain

This is a high-level overview of the different interactions between the Blockchain and other elements:

- All the NANCY devices that are either providers or consumers of resources will be authenticated and registered on the NANCY blockchain and hence will have access to the ledger.
- End-users will authenticate onto the NANCY blockchain by using the PQC means that TDIS will provide (smart card).
- The public part of the PQC credentials will be stored on the NANCY blockchain.
- The SSI infrastructure that will be proposed should have a data registry that keeps cryptographic material for the SSI credentials. The NANCY blockchain can be used as the data registry.
- All the NANCY devices that are either providers or consumers of resources will adopt one of the three roles as explained before (orderers, nodes or clients). For instance, an end-user device, once authenticated onto the Blockchain, will be able to consult prices and conditions as expressed in the Blockchain, by paging one of the nodes, and will be able to sign a transaction.
- The Marketplace is a set of providers and consumers that exchange goods (resources and services) while maintaining authenticity, integrity and privacy. This exchange must be performed, as in the non-digital world, through a contract, which in the case of NANCY will be a smart contract. A connection between the Marketplace and the Blockchain must therefore exist.
- During operation, the Smart Pricing module periodically updates the smart contract data. Thus, part of its logic must be connected to the Blockchain.

Interconnections and dependencies with other results

The main dependency is probably R5, which is the Post-Quantum Cryptography solution by partner TDIS, but other elements of the architecture are connected to the Blockchain in one way or another, as explained in the previous and following paragraphs.

PQC will reinforce security for devices (end-users) connected to the NANCY Blockchain. And the Blockchain shall be equipped with a means to check transaction signatures coming from those devices.

3.1.4. [R4] Realistic blockchain and attacks models and an experimental validated B-RAN theoretical framework

Overview of the technical advancements and innovations state-of-the-art at the date of writing

The foundation for the development of a secure, energy and spectral efficient wireless network, as well as the implementation of network optimizations, lies in the precise modeling of blockchain technology. In essence, blockchain plays a pivotal role in the establishment of a decentralized and self-governed B-RAN architecture. In their study, the authors of [4] examined and analyzed a range of trade-offs encompassing the ideal choice of block size in relation to throughput, as well as the delicate balance between increasing security through multiple confirmations and minimizing delay. However, despite the formulation of optimization problems, an analytical solution for minimizing latency was not offered, primarily because there was a dearth of suitable theoretical models for the blockchain. The experimental findings shown in the same work underscore the significance of considering factors such as power consumption constraints of user equipment and the bottleneck associated with blockchain scalability. Driven by the aforementioned issues, the authors in [84] introduced an initial queuing-based model to analytically assess the balance between latency and security. One limitation of this model is its failure to consider the impact of decentralization on RANs, alongside various crucial factors including block length, energy consumption, bandwidth availability, and other significant mechanisms such as the diverse types of consensus mechanisms applicable in B-RANs. Recently, researchers have made progress in developing non-analytical models for alternative and practical techniques. For instance, in [17], a method was proposed that utilized Fast Smart Contract Deployment (FSCD) and Digitized Spectrum Assets (DSA) in a B-RAN system. The aim of this method was to decrease access delay and enhance spectral management.

Purpose of the component redesign in view of GAP defined in 2.3

It becomes evident that current research lacks the ability to quantify several B-RAN aspects including: (a) the influence of decentralization on RANs following the implementation of blockchain technology; (b) the balance between the degree of decentralization and various KPIs such as latency, throughput, and energy consumption; (c) the various consensus mechanisms employed; (iv) the risk of alternative history attacks; and (d) the inherent limitations associated with blockchain technology. By developing novel theoretical and AI-based techniques for modelling the B-RAN and its attacks, this result will enable the quantification and estimation of the performance of B-RAN. Towards the same direction, various design choices will be taken into account and their performance will be evaluated. Performance insights and equilibriums will be revealed, while design and performance optimization guidelines will be highlighted. The aforementioned outcomes of [R4] will aid in filling the architecture, energy efficiency, and security gaps identified in section 2.3. Finally, the developed AI models will increase the intelligence of the NANCY ecosystem.

Technical advancements and innovations progress beyond the state-of-the-art at the time of writing

NANCY is driven by the aforementioned factors and aims to advance beyond the state-of-the-art by formulating a novel theoretical model that is experimentally validated. This model effectively incorporates the influence of decentralization in B-RANs, utilizing several mathematical frameworks such as Markov theory, random matrix theory, stochastic geometry, and machine learning techniques. This model will additionally consider various consensus techniques, sources of security vulnerabilities, potential attackers, and the inherent constraints of blockchain technology. It aims to identify the potential factors that have an influence on security and assess their respective impacts. Expanding upon the aforementioned model, the theoretical framework will be presented in order to assess

performance, specifically in relation to throughput, energy consumption, computational cost, latency, and security. This framework will consider the network parameters and the degree of decentralization as key factors. Furthermore, NANCY aims to establish KPIs that are special to blockchain technology, which will effectively measure the transparency, immutability, and availability of the B-RAN system. These KPIs will be carefully developed and subjected to rigorous analytical evaluation. Put simply, the theoretical framework is anticipated to uncover intrinsic trade-offs between the networks and blockchain-specific KPIs, which will be utilized to enhance the performance of B-RANs.

Contributions towards the realization of the NANCY architecture

The result at hand involves addressing the significant challenge of comprehending the specific characteristics of the NANCY B-RAN architecture and the various forms of attacks it may encounter. This understanding is essential for both the theoretical evaluation and the enhancement of network performance. In order to advance in this particular path, NANCY will utilize both experimental measurements and modelling data. In this respect, a Markov-based model is employed to incorporate the essential attributes of B-RAN, including the birth and death of requests and blockchains, with greater elaboration. The utilization of stochastic geometry and random matrix theory will be employed to create a model that represents the dynamic resource and storage capacities of the network, as well as the physical attributes of the end-nodes. These attributes include the number of end-nodes, their spatial positioning, and the resources and capabilities of each node to function as a connectivity provider. Furthermore, this endeavour involves the identification and quantitative modelling of various sorts of attacks. In order to offer a manageable model that may provide valuable design and optimization guidance for B-RAN, as well as uncover the associated security and privacy vulnerabilities, a machine learning-based technique will be employed. During the initial stage of the project, the ML-algorithm will undergo training using simulation results. Following the establishment of in-lab testbeds, the substitution of simulated data with experimental data will be undertaken to enhance the precision of the model. These constructed models will be utilized to quantitatively assess the attainable performance of the B-RAN in various environmental conditions. The evaluation of coverage probability and outage probability in an urban context will be conducted through the NANCY testbeds. The evaluation of performance at the network level will be conducted, and the resulting simulation and experimental outcomes obtained from NANCY will be presented and juxtaposed with the established theoretical framework. At the final stage of NANCY, the data obtained from the outdoor testbeds will be utilized to refine the models and validate their precision.

Interconnections and dependencies with the other results

The realistic blockchain and attacks models and experimental validated B-RAN theoretical framework receives the “Use Cases Descriptions, Network Requirements and Specifications as well as the Technology Enablers and KPIs Definitions” that will be delivered by D2.1. In addition, the two versions of the datasets developed in the Greek and Italian testbeds will provide the necessary information to perform the experimental evaluation of the models. This result will develop the theoretical framework that will provide insights into the performance that can be achieved by the B-RAN architecture [R1]. In addition to the analytical approaches, this result will utilize intelligent ones for estimating the performance of B-RAN. The developed models will be stored in the self-evolving AI model repository established within [R9]. Finally, the explainability of the intelligent models that will be developed in [R4] will be quantified through the explainable AI framework [R12].

3.1.5. [R5] A novel quantum safety mechanisms to boost end-user privacy

In quantum computing, the rules of quantum mechanics are exploited to process information in ways that are impossible on a conventional computer, thanks to the basic unit, the "qubit[2], whose particularity is that it is in a superposition of 2 states at the same time, instead of being in a single state. This particularity can be used to define quantum algorithms that take advantage of this superposition property in order to improve the execution time and then execute these algorithms on quantum processors.

Quantum computing would solve problems faster than conventional computing can do. For mathematical problems used in cryptography, quantum computing is expected to provide an exponential speed-up compared to conventional computing.

To bring answers to such quantum computing threats, 2 major tracks are envisaged by the community.

The first track being the Quantum Key Distribution (QKD), sometimes also called quantum cryptography, is a mechanism for agreeing on the encryption keys between remote parties, relying on the properties of quantum mechanics to ensure that the key has not been observed or tampered during transit.

The second track being the Post-Quantum Cryptography (PQC), is the area of cryptography in which systems are studied under the new quantum computing assumptions.

For some specific problems such as mathematical problems used in asymmetric cryptography, quantum computing is expected to provide an exponential speed-up compared to conventional computing. This means that the factorization of large composite numbers (on which RSA security is based) or the computation of discrete logarithm (on which DSA and ECDSA are based) would become feasible with the use of a quantum computer regardless of the sizes of the keys.

A conservative approach regarding post-quantum symmetric cryptography is to double the key size (i.e., migrating from AES-128 to AES-256) and increase the digest size (i.e., migrating from SHA-256 to SHA-384). However, it seems quite clear from experts [85] [86] [87] that the Grover algorithm (which could theoretically be used to weaken the security of block ciphers and hash functions) will provide little or no advantage for attacking symmetric cryptography or hash functions. This means current applications can continue to use AES with key sizes 128, 192, or 256 bits (and equivalent symmetric or hash algorithms such as SHA-256, SHA-384 or SM4).

A lot of organizations have started working on new asymmetric algorithms to replace RSA, DSA, ECDSA, DH. The most expected is NIST Project on PQC [88]. Third Round Candidate selection was announced in July 2022; these are the algorithms to be standardised (in grey the KEM algorithms for new Round 4):

Mechanisms	Name	Category	Note
Signature			
	Crystals-Dilithium	Lattice	Recommendation (strong security and excellent performance)
	Falcon	Lattice	For use cases where Dilithium signature is too large
	SPHINCS+	Hash	Based on another category

[2] Qubit stands for quantum bit - a qubit is the quantum state that represents the smallest quantum information storage unit and can be viewed as the quantum analogous of a bit

KEM			
	Crystals-Kyber	Lattice	Recommendation (strong security and excellent performance)
	BIKE	Code	Round 4 to be decided
	HQC	CodeLattice	Round 4 to be decided
	Classic McEliece	Code	Round 4 to be decided

Table 1 - NIST PQC Selection from Round 3

The targeted publication for the new standard on PQC signature and Public-Key Encryption/KEM algorithms is 2024.

[R5a] Quantum Key Distribution

Quantum communication is a subfield within the realm of quantum physics that investigates the transfer of quantum states or quantum information between one or more entities for specific objectives. Quantum communication lines and nodes, such as quantum repeaters and routers, collectively form what is known as a quantum network [89]. Quantum networks encompass a range of configurations, spanning from basic photonic devices with the capacity to manipulate and assess a single quantum bit (qubit) at a time, to extensive networked quantum computers. Quantum networks serve as the fundamental components of the prospective quantum internet [90]. One notable distinction between a quantum network and a classical network lies in the inability to accurately predict information exchange by extrapolation based on classical models. Within a quantum network, the process of obtaining information from a qubit necessitates a measurement. According to the principles of quantum physics, once a qubit is measured, it undergoes a collapse into a specific state, hence resulting in the loss of superposition and entanglement. Put simply, once a qubit is undergoing processing, it cannot be measured until a specific point in the calculation is reached, as dictated by the protocol governing the qubit. The various obstacles encountered have resulted in the predominant utilization of quantum communication in the realm of cryptography [91]. In this context, the principles of quantum mechanics are leveraged to safeguard data, particularly through the establishment of a shared secret key between mutually agreeable entities. This technique is commonly referred to as QKD and involves the transmission of a secret key across a quantum channel utilizing quantum particles, specifically photons. After the process of key sharing has been completed, the subsequent transmission of information is carried out exclusively over the classical channel. The categorization of quantum key distribution techniques is based on the detection technique necessary for retrieving the key information. Discrete-variable (DV) protocols and distributed phase reference (DPR) protocols are dependent on the utilization of information that is encoded on individual photons [92] . In the case of DV protocols, this information is encoded through the polarization of the photons, while in DPR protocols, it is encoded through either the phase or arrival timings of the photons. Consequently, both types of protocols necessitate the use of techniques for detecting single photons. Protocols involving continuous variables (CV) utilize coherent states to encode information about the quadrature of the quantized electromagnetic field. Consequently, homodyne or heterodyne detection techniques are employed in such scenarios.

Overview of the technical advancements and innovations state-of-the-art at the date of writing

While blockchain is often regarded as secure, it is susceptible to assaults from quantum computers [93]. Multiple research efforts have concentrated on post-quantum blockchain solutions designed to enhance the security of the blockchain using post-quantum cryptography [94] [95] [96]. However, QKD shows great potential in addressing the unique obstacles that blockchain technology will encounter in the quantum age. The viability of implementing a quantum-resistant blockchain platform using QKD for authentication has been shown in an urban QKD network [97]. In addition, a framework for a

permissioned blockchain that is protected using quantum technology and relies on a digital signature technique based on QKD has been introduced in [98]. Hence, the integration of QKD networks with blockchain technology to provide a robust and impregnable blockchain platform has emerged as a stimulating area of study.

Permissioned blockchain networks frequently handle substantial volumes of sensitive data. While the primary audience for this information may consist of other entities within the network, it is imperative to ensure the preservation of data confidentiality during its transmission. At present, the safeguarding of data secrecy relies on the implementation of conventional public-key cryptography systems. However, it is anticipated that these measures will be inadequate in countering potential eavesdropping threats posed by quantum-capable adversaries in the future.

QKD currently possesses a limited market share, although it is anticipated to experience exponential growth in the foreseeable future. The primary obstacles encountered thus far in the implementation of commercial QKD systems revolve around the intricate task of constructing robust and scalable devices capable of facilitating long-distance communication [99]. Due to this, it is advisable to maintain the connection length below 100 kilometers. Despite this limitation, QKD allows the circumvention of traditional secure repeaters for the purpose of storing and transmitting the keys between peers. This would indeed constitute not only a complication but also the most vulnerable aspect of the system in terms of security. The fronthaul emerges as the most suitable option for this trial due to its possession of all the necessary attributes and widespread architectural presence, hence enabling the optimization of benefits in terms of prices, power consumption, and dependability. The system operates effectively over short distances, necessitates the use of traffic encryption, utilizes coarse and dense wavelength division multiplexing (WDM) optical transport to enable a high density of logical channels on a single fiber, and minimizes the presence of active components (such as optical amplifiers) along the physical link to avoid compromising the integrity of the quantum signal.

Purpose of the component development in view of gaps defined in 2.3

The objective of this result is to explore the advancement of methods that utilize the principles of quantum mechanics to facilitate the secure sharing of a secret key between symmetric parties within the context of the B-RAN architecture. Specifically, innovative QKD mechanisms will be devised to effectively produce and exchange quantum keys using simulations and experiments. In this respect, [R5a] will focus on filling the architecture and security gaps identified in 2.3.

Technical advancements and innovations progress beyond the state-of-the-art at the time of writing

In order to demonstrate the simulation of a Quantum Link for exchanging private keys between components within NANCY architecture, TECNALIA´s Quantum Key Distribution laboratory services will support functional mockups based on NANCY use cases. The QKD laboratory has QKD commercial hardware made by the Switzerland company ID Quantique and it is working towards deploying the needed infrastructure to set up a 1 Km optical link between trusted nodes. The devices use time-bin encoding and Coherent One Way protocol in the physical layer.

This physical QKD deployment will set the ground to perform a wide range of communication experiments were two parts need to exchange information securely using symmetric cryptography with the benefit of the security against eavesdroppers implicitly associated with quantum mechanics.

Towards the realization of the NANCY architecture

To make this infrastructure available to NANCY use case, a REST-based Application Programming Interface (API) will be developed for both trusted nodes to be integrated in the same use case used and will facilitate the retrieval of quantum keys through user requests. These APIs will be compliant

with ETSI-014 standard for QKD private keys consumer to generate and request keys from the nodes and will be reachable via the internet in a secure way to be defined. The keys length generated by the QKD system can be 128 or 256 bits, depending on the use case definitions.

Once the agreed NANCY functional mockup use cases are tested against the simulated and the physical link, there will be a comparative report between the parameters recorded during the tests in both scenarios like KeyRate, QBER, and visibility to extract conclusions about the reliability of the simulations and the applicability of this technology in actual scenarios like NANCY.

The utilization of coarse and dense WDM techniques will be employed to facilitate the effective transmission of encrypted data and quantum keys.

Interconnections and dependencies with the other results

This result takes as input the use case requirements that will be derived from WP2 as well as the overall NANCY architecture designed in WP3. The QKD techniques that will be developed in this result will be incorporated into the B-RAN architecture [R1] for increasing point-to-point security in direct connectivity scenarios. In addition, the utilization of QKD in combination with blockchain technology will be investigated throughout [R3], [R4], and [R5a].

[R5b] Post-Quantum Cryptography for Digital Signature

Overview of the technical advancements and innovations state-of-the-art at the date of writing

All the major state agencies have provided strong recommendations enjoining all security networks and stakeholders to prepare and migrate to the quantum era on the horizon of 2030.

In the US, the White House mandated federal agencies to define post-quantum trajectory for critical systems by the summer 2022 and NSA plans a transition for National Security Systems for 2025-2033.

In Germany, the Federal Office for Information Security (BSI), has also defined migration recommendations towards the 2030 horizon with the use of **"**hybrid protocols" combining classical and quantum-resistant primitives, as this combination should protect both against conventional and quantum threats. BSI also calls for "crypto-agility", to make the cryptographic mechanisms able to react to all possible security events.

In France, the national security agency (ANSSI) calls for a three-step transition period to be concluded in 2030. Up to 2025, hybridization to provide some additional post-quantum defense-in-depth to the classic security assurance. During 2025-2030, hybridization to provide post-quantum security assurance while avoiding any pre-quantum security regression. From 2030, possibility to hand-over with only post-quantum cryptography.

Purpose of the component development in view of gaps defined in 2.3

The objective of this result is to provide to the NANCY framework the cryptographic primitives ensuring the security resilience of public key cryptography towards the coming quantum computing era. In particular, the new quantum-safe cryptographic algorithms will be implemented on a tamper-proof hardware device bringing, thus, the highest security protection to the framework.

This tamper-proof security device will be used to:

- Ensure authentication of B-RAN
- Provide a Quantum-safe Signature token for securing the communication

In this respect, [R5b] will focus on filling the Architecture and Security gaps identified in 2.3.

Technical advancements and innovations progress beyond the state-of-the-art at the time of writing

TDIS was participating in the ELECTRON H2020 Project [100] linked with energy management (EPES). During this project (from October 2021 to September 2024), TDIS focused on the development of the digital signature based on PQC into a smart token. As a project result, a first version of the PQC Hybrid Digital Signature Solution has been provided.

This initial Solution has validated the following innovations:

- Frugal implementation of PQC Digital Signature on tiny CPU devices environment (32bits CPU, 24kB RAM)
- Selected algorithm: Crystals Dilithium-AES, security level 2
- High secure implementation of the cryptographic algorithm including countermeasures against state-of-the-art attacks (side channel, fault injection attacks)
- Acceptable performance compared to classical cryptography
- Validation of a hybrid concept that consists of a combination of pre-quantum and post-quantum cryptographic algorithms

During this ELECTRON project, TDIS followed closely the PQC standardization process conducted by NIST, along with the recommendations from National Security Agencies (ENISA [101], ANSSI [102], BSI [103]) with the combined objectives of security and interoperability.

During the course of the ELECTRON project some new events occurred from the NIST standardization: adoption of a different variant Crystals Dilithium-SHAKE and recommendation to target security level above 2. This is the first reason explaining the need to redesign the component from the ELECTRON project.

Another event coming from National Security Agencies is the strong recommendation to implement a capability for Crypto Agility. Crypto Agility allows for a system or application to migrate to alternate cryptographic algorithms without causing a significant disruption to the infrastructure, allowing security updates to be quickly deployed to fix broken algorithms or replace vulnerable ones. In short, Crypto Agility offers the flexibility to meet the changing security needs of our connected world. This Crypto Agility is not supported in today's smart tokens due to their limited resources.

To cover the gaps described in the two above paragraphs, within the NANCY project, TDIS's objective is to work on the following innovations on the PQC Digital Signature component. Starting from the existing component, TDIS will develop the Crystals Dilithium SHAKE with Security Level 3 as recommended by NIST.

The new challenges for this innovation will be to keep the same requirement constraints while implementing a much stronger PQC algorithm:

- Frugal implementation of PQC Digital Signature on tiny CPU devices environment (32bits CPU, 24kB RAM)
- High secure implementation of the cryptographic algorithm including countermeasures against state-of-the-art attacks (side channel, fault injection attacks)
- Acceptable performance compared to classical cryptography

Moreover, TDIS targets to bring a novel mechanism for Crypto Agility as recommended by National Security Agencies. As of today, such capability is not supported on smart tokens.

Towards the realization of the NANCY architecture

Within NANCY architecture, TDIS will introduce a PQC digital signature in order to ensure the security resilience of Blockchain. Current Blockchain technology uses public-key cryptography (ie. like the ECDSA scheme) and hash function (ie. SHA) for signing transactions. The hash function SHA-256 is, today considered, quantum-safe. A conservative approach would be to increase digest size (e.g. move from SHA-256 to SHA-384) but, as of today, this is not considered a requirement within the NANCY project.

Public Key cryptography must be replaced with a quantum-resistant scheme. In order to ensure the security and integrity of Blockchain, Public-key cryptography is used to establish a distributed consensus of trust. While the chain itself is relatively secure, the wallets at the endpoints have already been demonstrated to be hackable, and quantum computing techniques will further expose them to security threats.

The solution to the blockchain wallet vulnerabilities problem is to create quantum-safe crypto wallets secured by PQC digital signature.

Interconnection with other results

To reach the objective of securing blockchain within NANCY architecture, TDIS PQC component will be connected to R3. PQC digital signature will reinforce security for devices connected to the NANCY blockchain. And the blockchain shall be equipped with a means to check transaction signatures coming from those devices.

[R5c] Post-Quantum Cryptography for secure communication

Overview of the technical advancements and innovations state-of-the-art at the date of writing

Despite the absence of quantum computers with the capability to break current encryption, security experts emphasize the importance of proactive planning due to the time required for integrating new algorithms into all computer systems. NIST has since released draft Federal Information Processing Standards (FIPS) for four chosen algorithms: CRYSTALS-Kyber for general encryption (FIPS 203), CRYSTALS-Dilithium for securing digital signatures (FIPS 204), SPHINCS+ for digital signatures (FIPS 205), and FALCON (anticipated draft FIPS in 2024) also for digital signatures.

Although these four algorithms constitute the initial post-quantum encryption standards, NIST is continuing its selection process for additional algorithms intended to supplement the initial set. The objective is to have backups based on distinct mathematical problems to provide alternative defense mechanisms in case vulnerabilities emerge in the selected algorithms. This necessity for backups was emphasized when an algorithm from the second set exhibited vulnerability after experts outside NIST successfully cracked SIKE with a conventional computer and we expect more to come.

The Open Quantum Safe (OQS) project is dedicated to advancing and testing cryptography that can withstand the challenges posed by quantum computing in an open-source setting. Within the OQS initiative, the primary focus lies in two key areas: first, the creation of liboqs, an open-source C library specifically designed for cryptographic algorithms that are resistant to quantum computing, and second, the incorporation of prototypes into a variety of protocols and applications, such as the widely used OpenSSL library. These tools not only aid our research efforts but also support the endeavors of other interested individuals or groups.

Purpose of the component development in view of gaps defined in 2.3

The PQC for secure communication component is a solution to address the security challenges presented in the gap analysis in Section 2.3. As quantum computers progress, the need for quantum-resistant cryptography becomes more urgent for ensuring the security of sensitive information.

Technical advancements and innovations progress beyond the state-of-the-art at the time of writing

The introduction of the PQC for secure communication components advances the PQC Digital Signature component through the following innovations. Building upon the existing OpenSSL framework and incorporating the OQS library. An added challenge in this innovation lies in the fact that the community is currently targeting only a limited number of operating systems and the work is rapidly evolving.

The results of this integration will be included in the application that has been prepared in order to measure the performances of the Ericsson 5G Edge (and compare it with the O-RAN solution). The application will integrate the results of the Open Quantum Safe (OQS) project. A traffic generator shall also be implemented to simulate a massive IoT scenario. The OQS library shall also be tested on a raspberry PI that will be connected to a 5G module. The server application shall also be tested for interoperability with the “Post-Quantum Cryptography for Digital Signature" from TDIS.

The library supports several new algorithms that will in the future be supported by the standard Openssl library. In order to test the algorithms before an official release is available the OQS ones are used instead and results are compared with standard algorithms. The target languages selected in this implementation are RUST and Python as they provide similar benchmark results.

Towards the realization of the NANCY architecture

In NANCY architecture the PQC component for secure communication will be integrated into the massive IoT testbed to guarantee a high level of security in the communication between IoT devices and the application running on the Edge server.

Interconnection with other results

The PQC for the secure communication component will be connected to PQC for Digital Signature component defined in R5b. The security of the communication will be improved using the digital signature module developed in the project.

3.1.6. [R6] Smart pricing policies

Technical advancements and innovations state-of-the-art at the date of writing

In terms of smart pricing policies, initially, in 2015, S. Hosny et al [104] proposed a mobile content marketplace that enabled mobile users to access and purchase various digital content such as music, videos, and e-books using their smartphones. Compared to other online marketplaces, this one offered a better user experience, as it provides a convenient and accessible platform for users to access content anytime and anywhere. The authors also proposed a pricing model that enables content providers to set different prices for their content based on various factors such as popularity, quality, and demand. Following this, Y. Jiao et al [105] proposed in 2018 a social welfare maximization auction mechanism for edge computing resource allocation in mobile blockchain networks. The authors argued that the proposed auction mechanism can effectively allocate computing resources to different users while maximizing the social welfare of the entire network. The proposed mechanism considered the heterogeneity of users' demands and computing resources and adjusted the auction parameters accordingly. The authors evaluated the performance of the proposed mechanism in simulations that showcased an increased performance compared to existing state-of-the-art mechanisms in terms of social welfare. Finally, K. Liu et al [106]. in 2019, proposed an optimal pricing mechanism for data markets in blockchain-enhanced Internet of Things (IoT) networks. The authors argued that the proposed mechanism can effectively incentivize IoT devices to contribute their data to the network while ensuring fair compensation for the data owners. The proposed mechanism used a combinatorial double auction model to match buyers and sellers and set the optimal prices for data transactions. The authors evaluated the performance of the proposed mechanism in simulations and showed that it can achieve high revenue for data sellers and low cost for data buyers while maintaining fairness and efficiency.

Purpose of the component development in view of gaps defined in 2.3

The Smart Pricing Framework is presented as a solution to address the challenges presented in the gap analysis in Section 2.3. Compared to SOTA, the smart pricing policies of NANCY need to adopt the security and privacy mechanisms developed within the project and most importantly to be scalable to the massive number of mobile users that NANCY envisions to involve in resource sharing and data relaying. Additionally, game theoretic pricing schemes should be considered to facilitate the balance between the strategy and choices of each user. The pricing policies of the NANCY project are strategically focused on παιempowering users to maximize their revenue through efficient resource sharing. This entails designing effective collaboration mechanisms among users within B-RAN (Blockchain-Based Radio Access Network), with an emphasis on monetizing the optimal allocation of network components and measuring the energy efficiency of user resources. By integrating these smart policies with a blockchain network, NANCY aims to bolster wireless networks by enhancing their reliability, durability, and sustainability, thereby providing a feasible alternative to traditional RAN (Radio Access Network) deployments. The application of AI techniques further contributes to the

project's goals by seamlessly connecting heterogeneous devices, reducing costs, and fostering a democratized and decentralized ecosystem, effectively addressing existing gaps in these critical fields.

Technical advancements and innovations progress beyond the state-of-the-art at the time of writing

In NANCY, Blockchain will be integrated into the RAN, this new addition will allow the development of the Smart Pricing Framework. Initially, the appropriate AI techniques will be sought in order to adapt and provide monetary user incentives and regulate resource sharing while retaining optimal profit models considering high revenues for data sellers and low costs for data buyers. Finally, the possibility of providing computational offloading incentives will also be considered. These can be discounted access fees and token rewards, for users who contribute their resources to the B-RAN through offloading. Reputation-based incentives for users who consistently contribute their resources to the B-RAN to earn a higher reputation score for them to access premium services or receive additional rewards and tiered pricing based on the number of contributed resources. In the context of the NANCY project, the integration of blockchain technology into RAN is poised to revolutionize the development of smart policies. To accomplish this, the project aims to identify and implement suitable AI techniques that can adapt and offer financial incentives to users while effectively regulating resource sharing. The overarching goal is to strike a balance between ensuring high revenues for data sellers and providing cost-effective solutions for data buyers. Additionally, NANCY envisions offering computational offloading incentives, which may include discounted access fees and token rewards for users who contribute their resources to the B-RAN through offloading. Reputation-based incentives will be considered as well, rewarding users who consistently contribute to the B-RAN with higher reputation scores, granting them access to premium services or extra rewards. Furthermore, tiered pricing models based on the quantity of contributed resources will also be explored, aiming to create a flexible and user-centric ecosystem for optimal resource allocation and profit models within NANCY.

Towards the realization of the NANCY architecture

We recommend that the Smart Pricing Framework be part of the Decision Engines, where it functions alongside the other components to determine the Access Point (AP) to which a User Equipment (UE) will migrate. Since NANCY will allow the UE to move between multiple providers APs, the Smart Pricing Framework is tasked with covering the monetary aspect of this decision. More specifically, the framework is presented with an array of available providers, and it employs game-theoretic methods, to identify the most optimal provider. The decision-making process revolves around determining the optimal price for the requested network resources that benefit every party involved in the transaction. These parties are the user’s main provider and the available new providers, that are eager to lease their equipment to attend to the user’s needs.

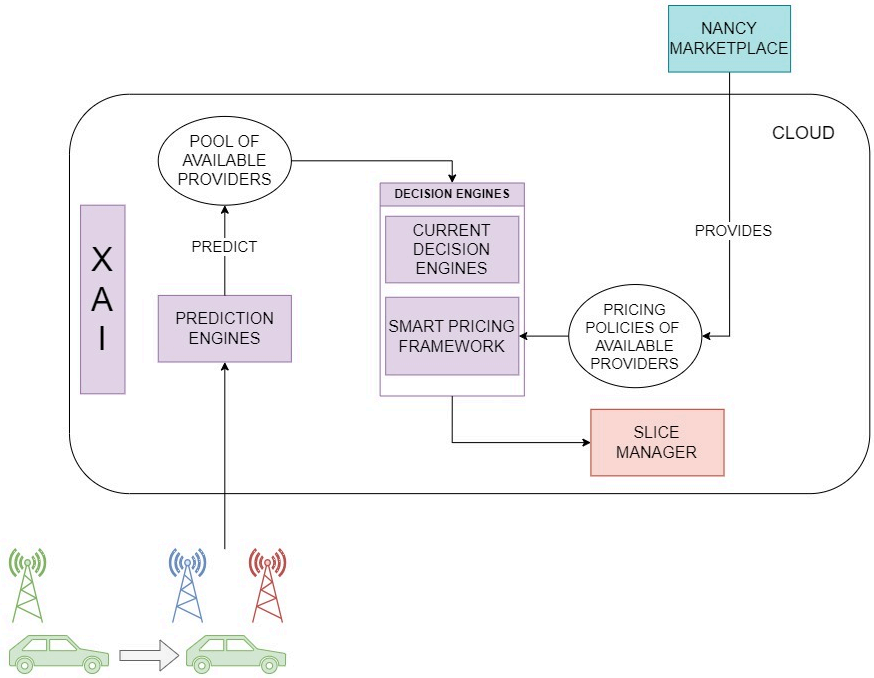

Figure 6 Smart Pricing Architecture

Interconnection with other results

The pricing policies within the NANCY project are intricately tied to several critical components and aspects of the system. Firstly, these policies are contingent upon the B-RAN architecture (R1), which forms the foundational structure of the network. Secondly, they are closely aligned with the grant and cell-free cooperative access mechanisms (R2), which dictate how users access and share network resources. Importantly, the pricing policies rely on network information pertaining to the routing and interaction with the Blockchain architecture (R3 & R4) to ensure the security and integrity of all network transactions, making them tamper-proof and resistant to unauthorized alterations. Additionally, the monetization of resources is strongly influenced by the availability and efficient allocation of computational and communication resources (R8). These resources are pivotal in driving the economic aspects of the system, as they enable users to participate and benefit from the network, making them a central component in the pricing policies' design and implementation.

3.2. Towards the Pareto-optimal AI-based wireless RAN orchestration that maximizes energy efficiency and trustworthiness

3.2.1. [R7] AI-based B-RAN orchestration with slicer instantiator

Technical advancements and innovations state-of-the-art at the date of writing

Orchestration in cellular networks is gaining remarkable attention due to the features that provide to coordinate different domains to manage the E2E lifecycle of services. These domains which normally are separated as Core, Edge, RAN, Transport, and Cloud take profit from inter-domain orchestration to achieve a seamless integration of resources of different categories (e.g. network, compute or radio). In this context, Network Slicing pools resources from different domains and interconnects them logically, isolating them from the rest of the resources, and performing dedicated allocation tailored to the services, user and/or network requirements [107]. A similar approach is also used to ensure the isolation of computing resources for software threads that implement radio functionality [108]. To this aim, lightweight orchestration approaches are considered as a key enabler, due to the shorter deployment time, ease of portability, and higher scalability. However, these lightweight shapes also bring challenges that need to be addressed, such as the need for container monitoring capabilities and the integration with specialized AI engines that provide self-healing and self-optimization capabilities [109]. The concatenation of different orchestrated VNFs forms a Service Function Chain (SFC), which is the base for deploying Network Slices. The SFC allows the independent management of each link in the chain, easing the scalability and adaptation to context changes, for instance, enhancing the service availability by configuring multiple paths to reach a service in high-demand scenarios [110]. Network slices are closely related, and normally used in conjunction with SFCs, Network Slices isolate the resources of different network segments, while SFC delivers a service inside the Network Slices. This also extends to CPU reservations, which can guarantee computing performance requirements in handling Network Slices for software threads implementing network functionality. This symbiosis is a perfect ground for AI engines to optimize the relation between the allocation, isolation, and performance of the services given certain requirements [111]. With the novel appearance of the O-RAN paradigm, as well as the emerging SDR technology, the orchestration of the RAN-slicing is paying relevant attention to AI working groups, which aim to optimize radio resource utilization [112].

AI can be employed for various tasks in O-RAN orchestration, including resource allocation and optimization [113]. While AI encompasses a variety of approaches and methodologies, reinforcement learning has emerged as one of the most promising solutions for fully automated slice instantiation that enables automatic perception of complex wireless networks, varying service requirements, and time-varying resource states through trial-and-error interactions with the environment. An example of such a solution was proposed in [114], where Deep Reinforcement Learning was used to develop an elastic resource reservation system for slices. The advantage of reinforcement learning lies in its ability to improve network resource utilization while minimizing potential violations of service level agreements compared to conventional rule-based resource allocation. Additionally, the authors in [115] demonstrate that reinforcement learning is a suitable approach for balancing user latency and energy consumption in a constrained environment for resource allocation. Reinforcement learning can also allow matching application-level, quality-of-service goals with resource-specific configuration parameters that are in turn used to enable fine-grained resource allocation using traditional approaches. Building on these advancements, we will develop an AI solution based on reinforcement learning for orchestrating B-RAN.

Purpose of the component development in view of gaps defined in 2.3

AI-based B-RAN orchestration is presented as a solution to address the challenges presented in the gap analysis in Section 2.3. The orchestrator introduces logical flexibility to NANCY architecture, providing a layer of abstraction that simplifies the integration of new technologies in different domains, avoiding the need to analyze the technical details of each underlying technology in each of these domains. Furthermore, the concept of resource orchestration between operators is proposed to achieve accommodating B5G networks and make B-RAN a sustainable alternative to traditional RAN deployments. Orchestration enables functions such as offloading and migration processes that are essential to maintain optimal communication service delivery, which is achieved by providing the necessary mechanisms to manage services at all stages: resource allocation, deployment, initialization, configuration, adaptation, and deletion. This functionality is especially valuable when it is required to interconnect resources in the form of Network Slices, enabling efficient and dynamic management of network resources. In addition, the envisioned orchestration engine is well integrated with the use of reinforcement learning and other AI techniques to manage radio, network, and compute resources autonomously in an optimized way, adapting to dynamic changes in the network, ensuring efficient network partitioning, and meeting specific application requirements. With this approach, the orchestrator also aims to address challenges related to the collection of unreliable data that could lead to misbehavior of the system, due to the capability of reinforcement learning engines to detect and ignore untrustworthy data.

Technical advancements and innovations progress beyond the state-of-the-art at the time of writing

AI-Based Orchestration capabilities will be leveraged by NANCY to perform autonomous management of radio, network, and computational resources as part of network slices. The management contemplates the coordination of different network segments to achieve E2E service delivery and deals with the interplay of requirements that interest both software workloads and radio traffic. Until now, AI-Based Orchestration has been applied to manage different and independent network segments, and at different layers. In NANCY we aim to pave the way to inter-operator resource orchestration providing mechanisms to enable resource sharing between different operators independently of the network segment and layer. But, in particular, radio resource orchestration in this context is of special interest, where Blockchain plays an enabling role, making trust mechanisms available for different purposes (e.g., Accounting, SLA assignment, Authentication...). Computing resources also need to be carefully managed to ensure that the orchestration goals mandated for radio resources can be adequately satisfied by the software threads that implement the radio functionality. This B-RAN orchestration is also envisioned to be driven by AI engines that control and enhance the orchestration procedures, such as optimization of radio resource allocation, placement, and VNFs configuration among others.

The use of reinforcement learning makes it possible to create an AI solution that can adapt to dynamic changes. Additionally, it is possible to take advantage of advanced techniques such as transfer learning and pre-train the solution using prior knowledge of the system based on predefined mathematical models. This serves as a starting point for the training process of the AI solution. In this way, we can develop a solution that can quickly adapt to the deployment environment and enable flexible and elastic network slicing. For example, upon application request, such as a video streaming application, the AI-based slicer would evaluate the specific requirements (e.g., bandwidth, latency, processing power, etc.) depending on the application, such as a video streaming application. It then selects the required resources to instantiate the network slice. In this scenario, the reinforcement learning solution would earn rewards for creating network slices that meet the requirements of the application without affecting the performance of other network slices. Over time, the agent would learn to instantiate network slices in a way that optimizes network performance.

Contributions towards the realization of the NANCY architecture

The result will aid the overall proposed NANCY architecture by helping to overcome the intelligence gap as identified in the Gap Analysis in Section 2.3. The selected approach of reinforcement learning allows the solution to learn from heterogeneous data. Additionally, the solution will be scalable and, as such, will be useful in large-scale deployments as well as smaller scales. Furthermore, reinforcement learning can be trained in an offline manner. This means that the system can respond in real-time, while training can occur without affecting the real-time response of the B-RAN. This capability ensures that the AI-driven solution can be continuously tested, refined, and updated, addressing the online training challenge that many AI approaches face such as in the example depicted in Figure 7. Lastly, as already briefly mentioned above, the reinforcement learning approach is robust and can handle real, unreliable input, even when it fluctuates over time. The latter directly addresses the third intelligence-related gap, ensuring that the designed AI-based B-RAN orchestrator will function optimally in various conditions.

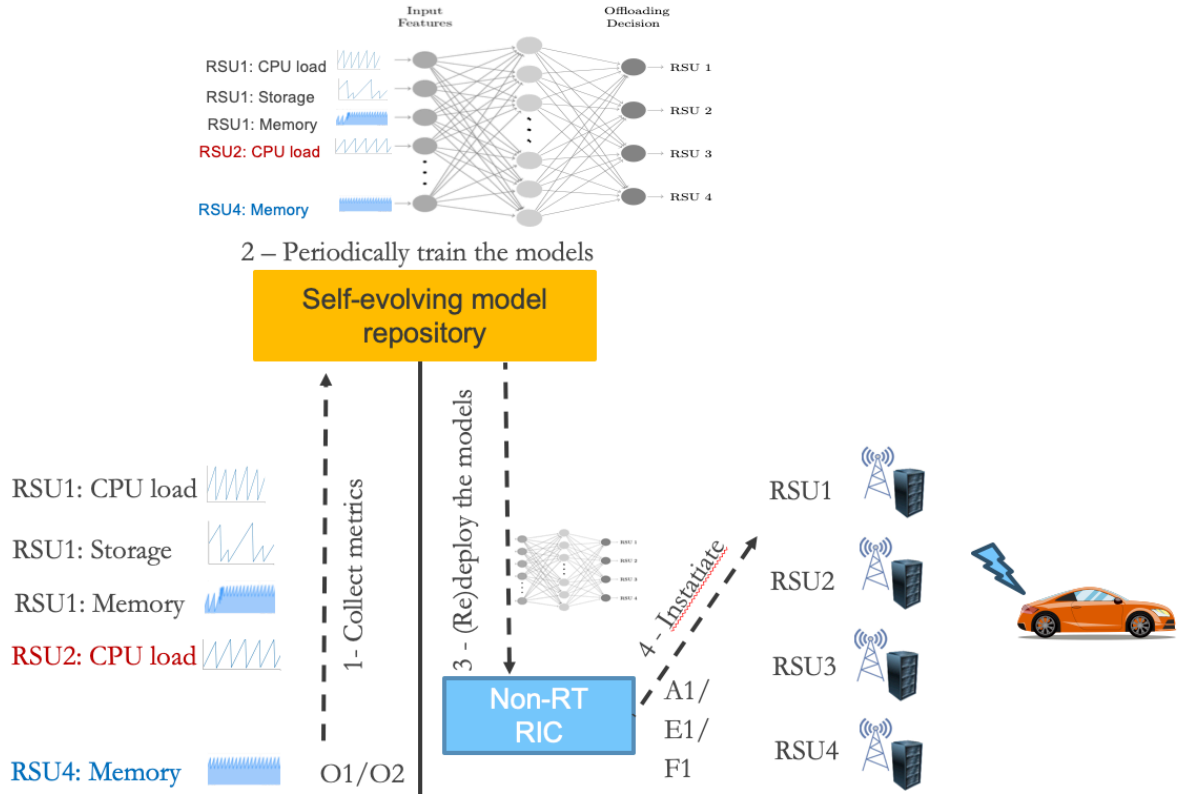

Figure 7 An example of integrating a self-evolving repository in the use case of offloading computation tasks to multiple roadside units.

Interconnections and dependencies with the other results

Since the AI-based B-RAN orchestration integrates a resource orchestrator between operators, it is interconnected with the B-RAN architecture [R1]. In addition, from [R9] Novel self-evolving AI model repository, it receives the latest version of the available deep reinforcement learning model, which enables dynamic model selection, continuous model improvement, and efficient management of diverse data for training and inference. It is also linked with [R10] Experimentally-driven reinforcement learning optimization of B-RAN, as this module attempts to improve efficiency and adaptability by focusing on problem-solving in intelligent components and addressing the challenges of data quality,

real-time decision-making, and adaptation to diverse network and resource conditions. Finally, AI-based B-RAN orchestration involves a key factor in achieving automation of task offloading to [R14].

3.2.2. [R8] A novel AI virtualiser for underutilized computational & communication resource exploitation

Technical advancements and innovations state-of-the-art at the date of writing

In the rapidly evolving landscape of telecommunications, the role of intelligent orchestrators has become pivotal. These orchestrators, powered by advanced technologies like artificial intelligence and machine learning, play a crucial role in optimizing network performance, enhancing user experiences, and streamlining operations. The European Telecommunications Standards Institute (ETSI) [116] presents two prominent examples of such intelligent orchestrators in Zero Touch Service Manager (ZSM) [117] and Experiential Network Intelligence (ENI) [118] working groups.

ZSM and ENI represent cutting-edge solutions at the forefront of intelligent orchestration in modern telecommunications. ZSM, with its advanced automation capabilities, is designed to streamline and simplify service deployment and management processes. Through a combination of machine learning algorithms and real-time data analysis, ZSM seeks to optimize network resources, ensuring that services are delivered efficiently and with minimal human intervention. On the other hand, ENI looks forward to leveraging the power of artificial intelligence and big data analytics to create a holistic view of network performance and user experiences. It utilizes intelligent algorithms to predict network issues, identify bottlenecks, and proactively address potential service degradation. Together, these intelligent orchestrators exemplify the future of telecommunications, where automation, data-driven decision-making, and predictive analytics converge to deliver seamless, high-quality services to users while optimizing network operations.

NANCY's primary objective with the AI virtualizer is to serve as a central component that could help the intelligent orchestration module. The fundamental responsibilities of the virtualizer encompass the identification of computational resources necessary for specific tasks and the subsequent facilitation of intelligent offloading decisions. To achieve this, NANCY strategically integrates several key technologies, including slicing, NG-SDN (Next-Generation Software-Defined Networking), NFV (Network Function Virtualization), and real-time CPU resource reservation. In summary, NANCY's AI virtualizer operates as an intelligent module that effectively identifies computational prerequisites and orchestrates resource allocation for both radio and computing resources. Leveraging cutting-edge technologies and intelligent ML algorithms, the AI Virtualizer enhances orchestrations' agility, thereby contributing to the overall efficiency and effectiveness of the orchestration process within the NANCY ecosystem.

Purpose of the component development in view of gaps defined in 2.3

In essence, O-RAN is designed to use ML to make networks more adaptive, efficient, and responsive to user needs, ultimately improving the quality and reliability of telecommunications services. Nevertheless, as of now, O-RAN has not introduced a fully operational intelligent orchestrator seamlessly integrated with ML technologies. The proposed AI-virtualizer within NANCY aims to enhance the existing orchestrator's functionality through the incorporation of ML interactions. These interactions seek to optimize resource utilization, harnessing available resources to their fullest potential, and enhancing the resource manager's adaptability and responsiveness.

Technical advancements and innovations progress beyond the state-of-the-art at the time of writing

The AI virtualizer is designed to perform intelligent resource orchestration in the Edge Cloud continuum by identifying the computational resources required by a specific task and making the

appropriate offloading decisions while enabling the exploitation of unutilized computational resources found throughout the NANCY edge-to-cloud continuum. This can be translated into two main tasks i) mitigate inter-slice conflict/contention when the aggregated resource allocation decisions exceed the available resources and ii) minimize underutilization by intelligently using the spare resources left by the concurrent slices without, however, exchanging any monitoring data between slices, which is essential to guarantee isolation and privacy. Given the non-scalability of centralized approaches, the envisioned AI virtualizer is based on a multi-agent deep reinforcement learning (MA-DRL) communication framework, where each slice is endowed with one agent responsible for resource orchestration via CPU scaling with dynamic adaption of reservation quotas and coordinates with fellow agents via a so-called communication action or message as shown in Figure 8. The DRL agents must cooperatively learn/discover the signaling policy, even without prior agreement on the meaning of control messages. They are guided by a reward function that penalizes conflicts and underutilization and minimizes latency. Fine-grained resource reservation with guaranteed CPU quotas and maximum service delay [108] will be employed to carefully control the utilization of computing resources.

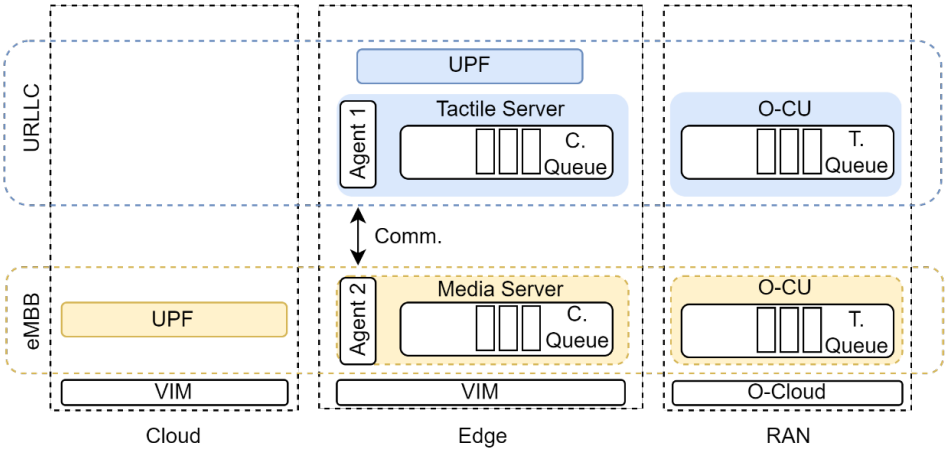

Figure 8 Multi-agent communication for inter-slice conflict and underutilization minimization.

Contributions towards the realization of the NANCY architecture

As depicted in Figure 9, the AI virtualizer minimizes inter-slice conflicts as well as CPU underutilization. This is due to its distributed architecture leveraging multi-agent communication, which consists of control messages rather than raw monitoring data. This contributes to reducing overhead and fostering scalability within O-RAN slices, bridging thereby the gap of current AI/ML deployment in O-RAN that requires a holistic view of the network training data.

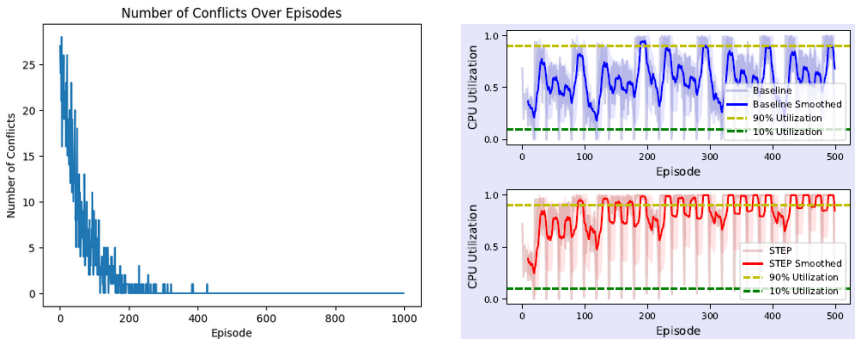

Figure 9 Conflict evolution vs episodes (left) and CPU utilization (right)

Interconnections and dependencies with other results

As NANCY's AI Virtualizer is a fundamental module in the orchestrator solution, it will actively interact with other state-of-the-art modules of NANCY. In particular, all the ML-OP and the Lifecycle management of the AI Virtualizer would be handled within the [R9] Novel self-evolving AI model repository. Additionally, [R10] Experimentally-driven reinforcement learning optimization of B-RAN will constantly ask the AI-virtualizer for the optimized version of the network.

3.2.3. [R9] Novel self-evolving AI model repository

Technical advancements and innovations state-of-the-art at the date of writing

Machine Learning Operations (MLOps), Artificial Intelligence Operations (AIOps) [119] are recent terms coined to focus on making ML model development and training more efficient. Such a need emerged particularly to democratize and industrialize AI as a service (AIaaS). While AI models and services are already scaled and well managed in cloud systems as recently shown with the plethora of large language models, reaching the same level of maturity for serving the needs of networks is still an open research and engineering endeavor [120]. Recently, such automation has also been investigated in the context of intelligent or AI native networks. One example is RLOps [121] focused on the development life cycle of reinforcement learning aided open RAN or for zero-touch service assurance in the H2020 5G-SOLUTIONS project [122].

Generally, evolving AI repositories are employed when adaptability and flexibility are the core goals of a 5G/6G network. PREDICT-6G [123], an H2020 project, deploys AI models to the control plane where predictability algorithms are required to proactively allocate 6G network resources. Since the management, and orchestration functionalities of the control plane are dynamic processes that adapt to the network's load, the AI models should be able to utilize new data to evolve and gradually produce better results. RISE-6G H2020 [124] goes beyond 5G networks and for this reason, AI algorithms are used to achieve intelligent, sustainable, and dynamically programmable services. 5G-IANA [125] focuses on Connected and Automated Mobility service provisioning over 5G networks. The project targets different virtualization technologies and leverages a Distributed AI/ML (DML) framework, as part of the virtual service repository, to provide end-to-end network services across different domains. 5G-CLARITY [126] supports network automation over multiple network slices by implementing AI-driven management capabilities. In this framework, the AI model repository interacts with the network through an SDN/NFV framework that provides a programmatic interface (API) for network configuration. The ARIADNE H2020 [127] project brings together a novel high-frequency B5G radio architecture with an AI network processing and management approach. This intelligent system mitigates the scale and complexity barriers of the new radio attributes, which cannot optimally operate using traditional network management approaches. In the domain of Edge AI, the AIatEDGE [128] H2020 adopts the serverless paradigm to provide a connect-compute fabric for creating and managing

resilient, elastic, and secure end-to-end slices. The envisioned architecture supports an AI model repository, which is also supported by hardware components, for resource management, service orchestration, and loop automation.

Purpose of the component development in view of gaps defined in 2.3

In addressing the complexities of AI-native RAN systems, introducing a self-evolving AI model repository emerges as a potential game-changer. Firstly, it tackles the issue of inference time by promoting a more efficient process: instead of relying on one large, monolithic AI model, it supports the utilization of smaller, nimble models, optimizing response times. This repository isn't just static; it offers dynamic selection capabilities, ensuring that the most adept or "expert" model is always in play for a given task or scenario. Beyond these, it also grapples with one of the most formidable challenges in the AI landscape: managing and making sense of heterogeneous data. This multifaceted approach promises to enhance the robustness and adaptability of AI-native RAN systems.

Section 2.3 identified four gaps, the relevant ones for this result are energy and intelligence. The selection of the model to be deployed impacts the energy consumption of the overall system as the deep neural network models are known to be resource-intensive. Selection strategies that take into account aspects such as target device, performance, and model complexity will be part of this result and contribute to the energy gap. With respect to the intelligence gap, the feature store will be able to manage the training features while the training and deployment pipeline will manage models and their deployment, resulting in an enterprise-like automation system.

Technical advancements and innovations progress beyond the state-of-the-art at the time of writing

The field of machine learning is vast and rapidly evolving, but the concept of a self-evolving model repository remains a relatively uncharted area in the research community. In the NANCY project, we aim to investigate and develop automated approaches to feature selection, model search, and selection based on ML/AIOps principles and tools [129]. Based on the NANCY network functionality reflected in the design of the APIs and the structure of the feature store [130], features can be selected, combined and engineered using classical feature interaction and enrichment approaches or representation learning [131] such as embeddings. Furthermore, the pipeline enables uploading and training contributed models on-demand or scheduled using neural architecture search and hyperparameter optimization techniques. This is also in line with recent findings [132] that suggest a trend toward favoring multiple specialized models over a singular, general-purpose AI model. Not only do these compact, highly specialized models require fewer resources, but they also offer faster inference times. Additionally, smaller models present fewer challenges in management, enhancement, retraining, and redefinition compared to their monolithic counterparts. This shift from the conventional monolithic model to dynamic model selection embodies a new era.

The model selection for continuous deployment is performed by an intelligent module that first identifies the relevant candidate models based on their (network) function and then selects the best based on a set of criteria through multi-objective optimization [133]. The candidate set selection can be realized through filtering through the semantic description of the model's function and I/O specifications, followed by dynamic model selection, though techniques such as Deep Reinforcement Learning and Mixture of Experts (MoE) approach [134].

Towards the realization of the NANCY architecture

The self-evolving AI model repository champions the idea of specialization, tailoring each model to excel in specific tasks or environments. This approach, in turn, provides a more streamlined, efficient path for AI model deployment and administration. By enabling fluid retraining and updating models

with diverse datasets and facilitating feature extraction from substantial data volumes, we aim to bolster the adaptability and versatility of the NANCY platform. More specifically, the NANCY self-evolving AI model repository is responsible for storing and searching for AI models by using low-complexity mathematical operations as its building blocks. It simultaneously searches based on the model, optimization procedure, and initialization parameters, thus limiting the human-design factor and highlighting the automated discovery of non-NN algorithms. This framework, as proposed in NANCY, represents every AI model as a computer algorithm with three (3) component functions, namely, setup, predict, and learn. These functions conduct the model initialization, prediction, and learning, respectively. To achieve this, NANCY will deploy a highly optimized open-source module that

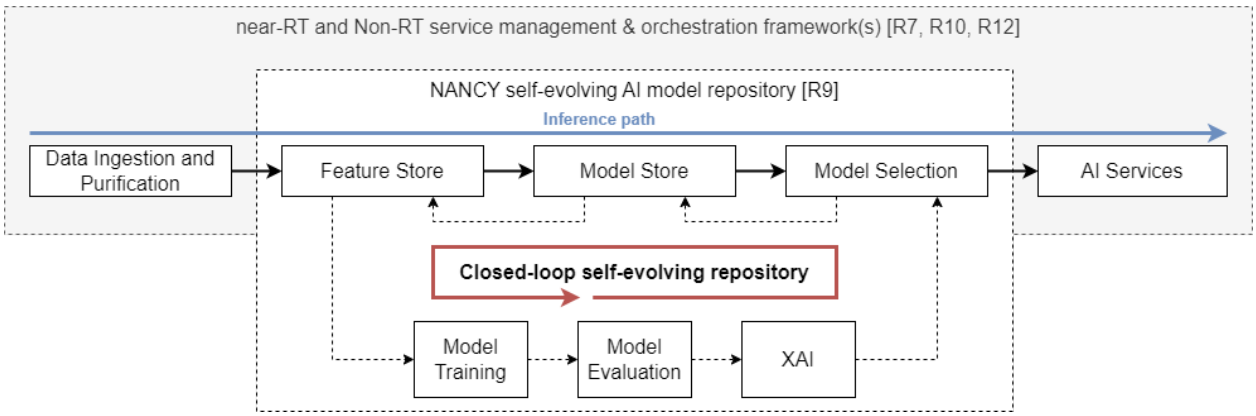

Figure 10 NANCY self-evolving model repository lifecycle

will be able to scan more than 10,000 models/second/CPU core.

Interconnections and dependencies with the other results

As presented in Figure 10, the NANCY self-evolving AI model repository will seamlessly integrate with both near-RL [R10, R12] and non-RL [R7, R12] service management and orchestration frameworks. Serving as a centralized hub, the repository will offer immediate access to a comprehensive collection of models to be used by R7 and R10. It not only supports optional dynamic model selection but also facilitates continuous retraining, updating, and enhancement of these models. Additionally, it efficiently manages a diverse range of heterogeneous data vital for both training and inference processes. For R12 services, a dedicated API will be available, enabling continuous monitoring of model performance, detection of spurious correlations, and guarding against potential adversarial threats.

3.2.4. [R10] Experimentally-driven reinforcement learning optimization of B-RAN

Technical advancements and innovations state-of-the-art at the date of writing

The total amount of available computation and storage resources in a B-RAN changes over time, in realistic scenarios. This raises the issue of resource and storage efficiency i.e., the optimal pro-active resource allocation scheme that enables the network to predict and utilize the optimized amount of resources at any time. Previous H2020 projects such as 5G-ZORRO [135] and Pledger [136] assumed a static model for resource allocation that enables the B-RAN to run in optimal conditions given a fixed amount of available resources. On the other hand, the dynamic nature of the network resources is considered in [137]. In this work, the authors propose a pro-active computational resource allocation model for C-RAN networks that increases the network QoS. Similar approaches have also been applied in SDN networks [138] and in fog computing environments [139]. Despite its success, this resource-aware mechanism has yet to be applied in decentralised B-RAN architectures. According to the existing literature [58], B-RAN is the ideal network to deploy such a technique due to its inherent resource flexibility and its dynamic operational environment.

Selecting the optimal technique for dynamic resource allocation is not a trivial task. Existing approaches utilize techniques such as deep reinforcement learning [140], weighted scheduling [141] and optimization algorithms [142]. Since a methodology, that would be considered optimal for all cases, is yet to be developed, most researchers focus on deploying application-specific approaches that satisfy the requirements of the system under examination.

Towards the realization of the NANCY architecture

Real-time intelligent components of the O-RAN require quick and accurate decision-making mechanisms that efficiently distribute available resources to interested parties. Therefore, resource allocation techniques should be redesigned to facilitate the following requirements: (i) Fast, real-time, and accurate decision-making to enable the deployment in realistic scenarios; (ii) operation with unreliable or noisy input to increase the generalization to a diverse set of deployment conditions; and (**iii**) adaptability to complex environments that contain high resource variability. Existing approaches fail to sufficiently deliver all of the aforementioned requirements and thus, they cannot be considered for deployment in real-world networks.

Technical advancements and innovations progress beyond the state-of-the-art at the time of writing

NANCY will develop a re-trainable optimization framework for resource allocation and will be empowered by reinforcement learning (RL) techniques. The consortium aims to implement a ML model to pro-actively allocate computation and storage resources over the deployed services. The framework will be trained with data that represent the network environment (e.g. how resource availability changes over time) and will employ an agent-based modeling approach to select the optimal amount of resources in real-time. The agent will engage in a decision-making process and for each decision, it will be rewarded through a reward function mechanism. Contrary to existing approaches, NANCY will employ a dynamic reward function that changes according to network conditions. This will enable the agent to converge faster to the optimal decision point. For example, in scenarios where high resource availability is observed, the reward function could generate higher paybacks and thus, the agent will aggressively favor new states. On the other hand, when low resource availability is observed, the reward function could yield lower rewards to decelerate the agent's state-hopping and influence its decision-making accordingly. Conceptually, this approach merges the learning rate of the neural network with the RL's reward function, achieving a faster convergence rate. As a result, it can be efficiently deployed in real-time systems that require quick decision-making.

Towards the realization of the NANCY architecture

Through the provisioning of the Experimental-driven reinforcement learning optimization of B-RAN, NANCY moves a step forward towards solving the existing issues with intelligent components in O-RAN networks. The proposed component will be designed to perform fast, real-time decision-making, and it will be tested in real operational environments. Special focus will be given to the fast inference operations where the models will be implemented to make decisions on the fly, achieving ultra-low latency delays. Further, NANCY will consider how noisy or incomplete data affect the ML models' performance and will employ feature extraction techniques to alleviate penalties stemming from such occurrences. Feature extraction, if properly designed, can help the ML models generalise better even when the data inputs are not of high quality. Towards this end, the proposed module will first expand the dimensionality of the input data through multiple layers of feature extraction operations and then, it will feed the extracted features to the ML model. In this way, data gaps or inconsistencies will be considered low-quality features and will play a trivial role in the decision-making process. In order to increase the model's adaptability to complex environments, NANCY will consider early-exit-inference mechanisms. Early-exit is a technique that enables the AI model to stop the inference process, according to a set of requirements and constrains. Such requirements would resemble the network's conditions such as the available resource types, or the functional requirements such as maximum latency. Since lower resource variability environments would not require the ML model to compute the inference operation until its very end, early-exit will terminate this process to generate predictions earlier. On the contrary, more diverse resource ecosystems would require a more complex inference process, in which case the early-exit mechanism will be disabled.

Interconnections and dependencies with the other results

The Experimental-driven reinforcement learning optimization of the B-RAN module is interconnected with the following results: [R1] *B-RAN architecture*, [R8] *A novel AI virtualizer for underutilized computational & communication resource exploitation* and [R9] *Novel self-evolving AI model repository*. B-RAN architecture feeds to [R10] since it provides the network organization details (number of nodes and B-RAN parameters). Such parameters are necessary for the AI model to describe the network environment where the agent acts upon. [R8] provides the resource utilisation rate of the network in real-time, which is important for defining the model states and the reward function. [R10] will periodically request updates from the [R8] in order to keep its internal state in synchronization with the network conditions. Finally, [R9] will be used to acquire the latest version of the DRL model available in the repository. Since the self-evolving model repository will engage with model retraining actions, in order to improve the performance of the stored models, [R10] is charged with fetching the corresponding model.

3.2.5. [R11] Semantic & goal-oriented communication schemes for beyond Shannon excellence

Overview of the technical advancements and innovations state-of-the-art at the date of writing

The inspiration for the idea of the 6G wireless networks, which encompasses semantics and effectiveness features, can be traced back to Nikola Tesla. In 1926, Tesla made a notable statement envisioning a future where wireless technology will be flawlessly implemented, resulting in the transformation of the entire Earth into a vast cognitive entity [143]. In this context, placing emphasis on semantics and accurately establishing and clarifying the objective of communication aids in extracting the data that are directly pertinent to effectively deliver the intended information from the source or achieve a predetermined objective. The implementation of a primary technique involves the disregard of unnecessary data, resulting in a substantial reduction in the volume of data that needs to be communicated and recovered. This reduction in data transmission and recovery leads to savings in terms of bandwidth, latency, and energy conservation. According to this, it is anticipated that goal-oriented and semantic communications would play a crucial role in investigating the significance of data and facilitating brain-like cognitive processes and efficient task completion among dispersed network nodes/entities. This shift in perspective signifies a significant change in paradigm, wherein the primary focus is on the successful completion of tasks at the intended destination (effectiveness problem), rather than solely emphasizing error-free communication at the symbolic level (technical challenge).

Historically, past iterations of wireless networks have been developed with the primary objective of accommodating the exponential increase in downlink traffic. However, a shift in the balance between uplink and downlink traffic has been observed since the fourth generation (4G) [144], resulting in a decrease and even a reversal of the previously existing asymmetry. The proliferation of algorithms in the context of 5G technology leads to a significant escalation of uplink traffic. Regrettably, the uplink capacity of 5G has not been adequately dimensioned to accommodate the rapidly increasing demand anticipated in the coming decade. 6G is anticipated to amplify this phenomenon as a result of the integration of a rapidly growing quantity of dispersed intelligent nodes that will gather, analyze, and retain data. The utilization of B5G technology in various sectors, such as industrial IoT and virtual reality, has led to the establishment of new KPIs. These KPIs encompass stringent latency bounds, packet delivery jitter, reliability, achievable throughput, and system dependability [145]. An instance of this can be seen in the case of holographic communications that utilize multiple-view cameras. It is anticipated that these communications will necessitate a substantial amount of data transfer, specifically several terabits-per-second (Tbps) per link, in both the uplink and downlink directions. This demand exceeds the capabilities of the 5G network. Additionally, there is a need for strict E2E latency to guarantee a virtual and seamless remote experience that closely resembles reality [146] [147].

Purpose of the component development in view of gaps defined in 2.3

In line with Tesla's conceptual framework, NANCY's approach to 6G networks prioritizes the incorporation of semantics and efficacy as fundamental components of NANCY's network architecture. The primary function of the NANCY semantic encoder is to identify and extract the semantic content from the source signal, thus eliminating any extraneous information. Consequently, the reduced amount of data that need to be transmitted through the network will improve not only data efficiency but also energy efficiency of the NANCY system. The NANCY semantic decoder is responsible for interpreting the transmitted information and converting it into a comprehensible format for the receiving node or entity. Additionally, it is important to evaluate the level of satisfaction of the receiving node in order to determine the extent to which the reception of semantic information can

be deemed successful. Ultimately, the objective of NANCY is to alleviate the presence of semantic noise, a phenomenon that arises throughout the communication process and leads to misunderstandings and erroneous interpretations of semantic information. The envisioned semantic communication approaches will contribute towards imbuing next-generation networks with native intelligence capabilities.

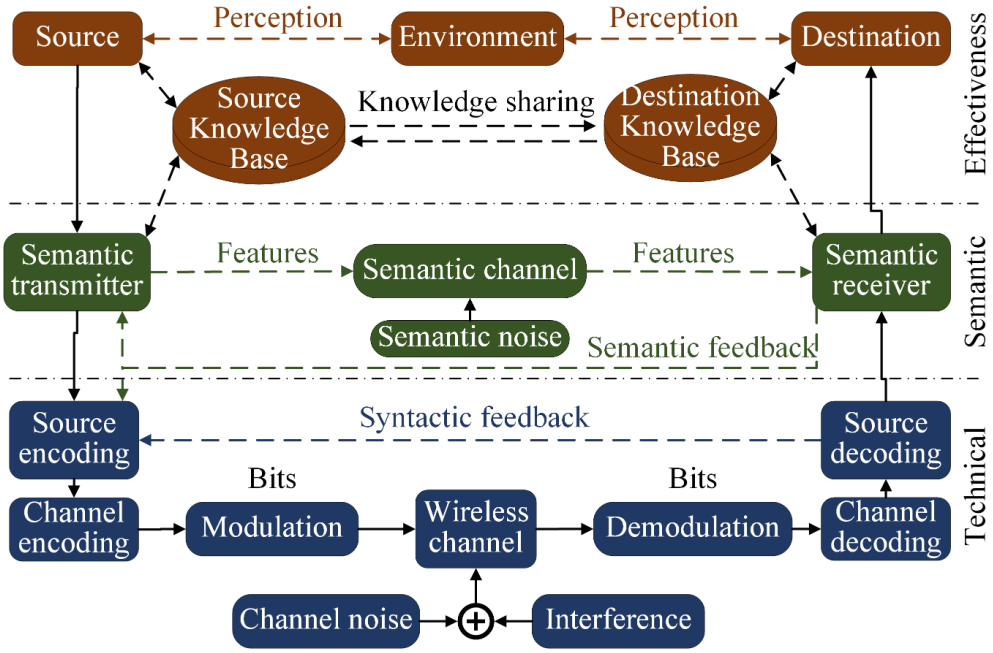

Figure 11 Semantic communications architecture

Technical advancements and innovations progress beyond the state-of-the-art at the time of writing

During the past five years, extensive research has been carried out in semantic communications, including textual data [148, 149, 150], voice signals [151], visual pictures [152, 153, 154, 155, 156], and video material [157]. The researchers in [158] presented a SemComs framework known as DeepSC, which utilizes the transformer concept. In the work done in [151], the scope was broadened to include voice transmission by using CNNs. In addition, the authors of [159] introduced a new neural network that integrates source and channel coding techniques to compress images. This network additionally utilizes CNNs to condense the given pictures. The researchers used a coding strategy that led to a 3 dB improvement in peak signal-to-noise ratio compared to traditional coding schemes in the presence of Rayleigh fading. The researchers in [160] devised an image semantic coding approach that employs the Laplacian pyramid to improve the picture compression ratio. The paper introduces a new method for adaptive picture semantic coding, described in [161]. This method employs reinforcement learning to improve the rate-perception-distortion metric, which measures the quality of image reconstruction. The authors of [162] conducted a thorough assessment of existing technical advances in semantic communications for intelligent wireless networks, prompted by improvements in adaptive image semantic coding. The assessment focuses specifically on the architectural elements, relationships between different layers, and diverse applications. This essay explores the difficulties that arise when trying to include semantic communications within the framework of the developing 6G wireless technology. Trying to include semantic communications within the framework of the developing 6G wireless technology.

Contributions towards the realization of the NANCY architecture

The objective of this result is to establish part of the communication strategies that will be employed in the project. The utilization of semantic communications will be employed to facilitate accurate transmission and understanding of information. In order to achieve this objective, an examination and adaptation of semantics and knowledge representation systems will be conducted to align with the specific demands of the B5G network. In addition, the communication semantics will be included in goal-oriented communications, wherein the exact limitations and specifications of each communication type and the connections between devices will be thoroughly delineated. Goal-oriented communications include transmitting only the necessary information to achieve a specific objective. This approach reduces communication overhead and enhances the energy efficiency of the system.

Interconnections and dependencies with other results

The semantic communication approaches that will be developed in this result will find application in various usage scenarios and use cases, with a specific focus on the ones analysed within the NANCY B-RAN architecture [R1]. In more detail, such techniques are expected to influence the B-RAN modelling and performance assessment carried out within [R4]. Finally, the models developed in this result will be stored and maintained by the self-evolving AI model repository [R9], while their explainability will be ensured through the explainable AI framework [R12].

3.2.6. [R12] An explainable AI framework

Technical advancements and innovations progress beyond the state-of-the-art at the time of writing

Radio Access Networks (RANs) involve numerous tasks that can be effectively handled using AI, such as management (energy, handover), detection of anomalies, and resource allocation among others. However, one of the main challenges faced in AI-driven RANs is the inherent lack of transparency in the decision-making processes governing these algorithms. Consequently, there is a need to integrate Explainable AI (XAI) into RAN AI operations. This integration would provide more comprehensive and detailed information about the decision-making processes of the algorithms, enabling better understanding and fostering trust in their operations.

Given the task under consideration, the availability of data, and the AI algorithm itself, various XAI techniques could be utilized to provide interpretability. In most cases, it is preferable to choose model-agnostic XAI techniques due to the fact that they can provide explanations without the need for a specific model. Forming tasks like management or anomaly detection as supervised machine learning tasks, techniques like LIME and SHAP could offer both local and global interpretability. Local Interpretable Model-Agnostic Explanations (LIME) [163] is a model-agnostic approach designed to tackle various data types such as tabular, and text. By creating locally linear models, LIME approximates complex learning models, effectively providing transparent explanations for individual predictions that would otherwise be considered a black box. On the other hand, SHapley Additive exPlanations (SHAP) [164] take a different approach by relying on feature relevance explanation to interpret specific predictions. They calculate an additive feature importance score while considering a predefined set of required properties. By utilizing the Shapley values method, SHAP determines the influence of each feature by estimating its marginal contribution to the final reward function. This enables a comprehensive understanding of how individual features contribute to the model's predictions. Another well-known model-agnostic technique is DeepLIFT [165]. DeepLIFT tries to provide feature relevance ratings for specific input characteristics, giving details on how each item affects the predictions that were made by the model. By describing ideas in terms of concepts rather than numerical quantities, Knowledge Graphs [166] provide a method for creating explanations that are

understandable by humans. The relationship between these ideas creates a knowledge graph, which turns out to be an effective method for data representation. It is possible to automatically create knowledge graphs, which may then be perused to learn new information about the subject area, including implied notions that weren't expressly declared. Since each step can be traced back, full explainability is enabled by this traceability, which makes Knowledge Graphs an effective tool for comprehending complex data better.

For the needs of XAI for resource allocation models, the Programmatically Interpretable Reinforcement Learning framework (PIRL) [167] could be used. PIRL utilizes a high-level, human-readable programming language to represent its policies. The framework employs Neurally Directed Program Search (NDPS), which enhances the interpretability of the generated policies, making them easier to understand and analyze. The Hierarchical Policies method [168] stands out due to its unique approach to the task description. In this method, each task is accompanied by human instruction, and agents can solely access learned skills through these descriptions. As a result, the policies and decisions derived from this approach are inherently human-interpretable, enhancing transparency and comprehensibility. Linear Model U-Trees (LMUTs) [169] were used also as a mimic reinforcement learning framework based on stochastic gradient descent trying to transform the RL problems into supervised ones, which are much more interpretable.

Purpose of the component development in view of gaps defined in 2.3

Deep AI models have recently gained high popularity due to their significant performance in various domains and tasks. However, their increased complexity has raised concerns regarding the way that decisions and predictions are being made, or how they handle sensitive data, which are currently treated as a black box. In particular, while certain AI methods, such as Decision Trees, can be treated as self-explanatory, Deep AI Neural Networks require a set of methods that will produce explanations regarding the decisions that were made.

This set of methods is under the umbrella of the eXplainable Artificial intelligence (XAI) which aims to provide a number of tools and techniques to any person that would like to analyse and investigate how certain decisions were made by Deep AI models. As a result, any individual can trust Deep AI models with high confidence, as they will no longer be treated as black boxes, while, simultaneously, any concerns about the privacy of data will be mostly eliminated as the whole procedure will be transparent.

In the context of NANCY, XAI techniques will mostly cover security and transparency aspects. In particular, Deep AI models will be mainly utilised for the optimisation of resource orchestration procedures; a set of processes that require high transparency regarding the rationale behind the best possible allocation of the resources that are available at a given time, placing network operators in a position in which they can insightfully explain the decisions that were made by the AI models. Furthermore, XAI tools can help in the identification of various attacks against the network, as they can provide insights regarding the incoming traffic, which can possibly result in the detection of an intruder in the network.

Purpose of the component development in view of gaps defined in 2.3

In the endeavor to unveil the inner decision-making mechanisms of black-box models, NANCY takes a step further by proposing a model-agnostic algorithm to make opaque models' decisions explainable and comprehensible. More precisely, NANCY proposes a 4-stage XAI engine, leveraging 2 of the most common key technologies for XAI, SHAP, and LIME. LIME [163] explains the model's behavior in a local vicinity by constructing sparse linear models around selected predictions. On the other hand, SHAP [164] provides global insights into the importance of specific features by estimating their average marginal contribution over all possible coalitions. In the first stage, the SHAP algorithm will be applied to extract the SHAP values, utilizing the corresponding Shapley values method. Then, these values will be used for producing descriptive elements like feature dependencies, explanations, and models' summarization among others. A feature selection process will be applied, promoting only specific features to the next stage, in which the LIME technique will produce local explanations, feature importance, and model evaluation. Finally, the outputs of the second and third stages will be fed into the final stage which will perform the diagnosis and decision-making tasks.

Contributions towards the realization of the NANCY architecture

XAI module will be an external component of the NANCY's architecture that will directly communicate through the most appropriate interface with the corresponding modules. In particular, the XAI module will mainly communicate with the non-RT RIC which is part of the System Management & Orchestration (SMO), and the repository which stores all the AI models that are utilized in the context of NANCY. Furthermore, it is worth noting that the explanations will be provided in a non-real-time manner, and they will be accessible to any interested individual with the required authorization at any given time. Finally, it is important to mention that the integration of the XAI module inside the non-RT RIC would be extremely challenging due to the nature of RIC, and, simultaneously, it would increase the complexity both of the architecture and the RICs individually.

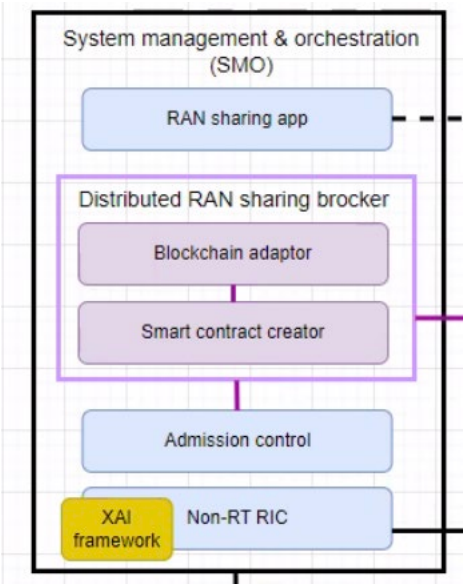

Figure 12 XAI as part of NANCY Architecture

Interconnections and dependencies with other results

XAI framework has a strong interconnection and is directly linked with "*R9: Novel self-evolving AI model Repository".* More specifically, regardless of the secure communication between the AI model repository of NANCY that is required, the XAI module should be able to support the interpretation and the explainability of the decisions and the predictions that novel AI models made. In order to achieve this, model-agnostic XAI tools and techniques will be mostly investigated and utilized focused on their level of adaptability.

3.3. Distributed MEC for "almost-zero latency" and high-computational capabilities at the edge, where the data are generated

3.3.1. [R13] Next-generation SDN-enabled MEC for autonomous anomaly detection, self-healing and self-recovery

Technical advancements and innovations state-of-the-art at the date of writing

Self-healing technology aims to maintain wireless cellular networks during the maintenance phase, and it becomes more crucial in 5G and 6G networks due to their highly complex and heterogeneous nature. These networks are also prone to faults and failures, with the most critical domain for fault management being the Radio Access Network (RAN). Each network element serves a specific area, making redundancy a challenging task. If a network element fails to fulfill its responsibilities, it results in degraded performance, and users suffer from poor service quality, resulting in a considerable revenue loss for the operator. To address this challenge, self-healing technology has been introduced in emerging cellular networks. It focuses on the automatic detection of network faults and failures and the implementation of the required corrective actions to mitigate the service degradation effect. The main defined use cases are the following:

- Cell Outage Management (COM)
    - Cell Outage Detection (COD): This function automatically identifies cell outages by using input parameters such as Key Performance Indicators (KPIs), alarms, and measurements. When the values of these parameters meet the COD condition, the cell outage is detected. For example, if the value of a KPI exceeds a predefined threshold or an alarm is triggered during a cell outage, the COD function identifies the outage.
    - Cell Outage Compensation (COC): This function automatically compensates for a cell outage to maintain cell operations. When a neighboring cell detects a fault, it classifies the type of fault and makes a compensation decision. The compensation can be in the form of relay-assisted handover (HO), power compensation, or reconfiguration of their antenna tilt. These compensation techniques help to mitigate the impact of the outage on the network and maintain the quality of service for the users.

The classes of problems that need to be addressed when managing the network autonomously are:

- Anomaly detection: This is a method commonly used in the diagnosis of network faults or misbehaviors, as it is designed to identify and flag abnormal behavior within the network. This approach relies on unsupervised learning methods and is particularly useful when it comes to detecting issues related to faults or improper network settings. Popular algorithms for anomaly detection include k-NN, Local Outlier Factor, and DBSCAN algorithm that is used in [170].
- Pattern identification: Pattern identification is a task that involves identifying patterns or groupings of cells that exhibit similar behavior, such as increased traffic or dropped calls. This kind of problem relates to COD issues and solutions can be found in UL and SL literature. For instance, a Decision Tree is an algorithm that can be used to classify cells into different groups based on their characteristics, such as traffic load and signal strength [171].
- Control optimization: This category of problems arises frequently in the realm of autonomous network management, where control decision problems are encountered to adjust network parameters online, with the goal of achieving specific performance targets. These types of decision problems, where the optimal decision is learned online based on the feedback from

the environment to the network's actions, can be tackled using RL [172], [173], and DRL techniques [174], [175]. COC problems can be solved using these methods.

Among the Supervised Learning techniques, we focus on Neural Networks (NN), which are a class of machine learning algorithms inspired by the structure and function of the human brain. They consist of interconnected nodes, or neurons, that work together to learn patterns and relationships in data. There are several types of neural networks, each with its own architecture and strengths. Currently, the application of neural networks in cell outage management is limited, however, an example of the use of a feedforward neural network (FFNN) in detecting cell outages is being reported in the literature [176]. However, Long Short Term Memory (LSTM), which is a type of recurrent neural network architecture that is well-suited for modeling sequential data, is particularly effective in capturing long-term dependencies and can be used to predict outcomes based on past events and thus for COD tasks. In the context of self-healing, article [177] utilizes this model for outage detection.

Among the Unsupervised Learning techniques, Clustering (and in particular K-means) and Outlier Detection techniques have been taken into account. K-means is used to group data points into k distinct clusters based on similarity. It works by iteratively assigning data points to their nearest centroid and moving the centroid to the mean of the assigned data points. K-means is efficient but sensitive to the initial placement of centroids and determining the optimal number of clusters. In self-healing, this algorithm is mainly used to detect cell outages [175], [174]. The article [175] focuses solely on users who are within range of the base station that has been interrupted, as well as its neighboring base stations. As for Outlier Detection, Local Outlier Factor (LOF) is one of the popular unsupervised outlier detection algorithms. LOF measures the local density of a data point compared to its neighbors and identifies points that have a significantly lower density as outliers. LOF is widely used in fraud detection, intrusion detection, and industrial quality control, among its applications in cell outage detection [172], [178]

As per Reinforcement Learning techniques (and in particular their deep variants, including Actor-Critic), we focused on Deep Q-Network (DQN), which is a DRL algorithm that combines deep neural networks with Q-learning. DQN has been used to solve various problems, including game playing, robotics, and natural language processing. The algorithm works by using a deep neural network to estimate the Q-values of each possible action given the current state. The network is trained using a variant of Q-learning called experience replay, which involves storing the agent's experience in a replay memory and sampling a batch of experiences randomly to update the network's parameters. DQN is known for its ability to learn directly from raw sensory inputs, making it useful for problems where hand-crafted features are not available or are difficult to design. In the [175] the author has applied DQN to solve the COC problem.

The Actor-Critic (AC) algorithm combines elements of both value-based and policy-based approaches. In AC, the agent maintains both a value function and a policy and uses these to guide its behavior in the environment. The value function estimates the expected long-term reward for a given state-action pair, and the policy specifies the probability distribution over actions for a given state. The AC algorithm consists of two components: the actor, which learns the policy, and the critic, which learns the value function. The critic updates its estimates of the value function based on the difference between the observed reward and the expected value, using the temporal difference (TD) error.

A summary of the self-healing techniques using ML techniques has been reported in Table 2.

Table 2 - A summary of the self-healing techniques using ML techniques.

	Reference - Year	ML technique	Objective	Algorithm
COD	[171] - 2015	SL/UL	Anomaly detection	k-NN, LOF
	[176] - 2015	SL	Diagnosis	FFNN
	[178] - 2018	UL	Pattern identification	LOF
	[170] - 2018	SL/UL	Anomaly detection	DBSCAN algorithm
	[177] - 2019	SL	Pattern identification	LSTM
	[176] - 2021	SL	Anomaly prediction	Prophet Algorithm
	[171] - 2021	SL	Pattern identification	Decision tree
COC	[173] - 2014	RL	Control Optimization	Actor-Critic
	[172] - 2015	RL	Control Optimization	Actor-Critic
	[175]/ [174] - 2020/2019	DRL	Control Optimization	K-means and DQN

Purpose of the component development in view of gaps defined in 2.3

Anomaly-detection, self-healing, and self-recovery address the challenges in the gap analysis presented in Section 2.3. Indeed, it ensures the services for the end-users to be reliable even in case of RAN faults (e.g., at gNB level) and/or at the MEC level (e.g., in case of attacks), which can be beneficial for the end users and for the whole NANCY architecture components.

As for security and privacy, the literature review showed that many papers in the literature make use of ML tools to provide self-healing services to the radio access network. However, no particular paper focuses on the privacy of users, for example using Federated Learning techniques (which require a central entity to store the complete model, which can be identified in the MEC of the RAN). Moreover, we plan to assess the possible improvements given by the decentralization of the Federated Learning techniques we will employ for the self-healing component, by using consensus-based techniques.

As for energy efficiency, anomaly-detection, self-healing, and self-recovery mechanisms based on AI and Federated Learning can reduce the energy consumption of the mobile network, by reducing (in the deployment phase, so after the training phase) the number of communications among FedL agents (indeed each agent has its own prediction model) and also the computational complexity is reduced in the operation phase, due to the constant/linear computational cost of the Neural Network prediction. Finally, energy efficiency can be achieved also at the RAN/MEC level, by optimizing resources in case of faults.

The advantage brought by AI methodologies of the proposed anomaly-detection, self-healing and self-recovery mechanism, and in particular Federated Leaning techniques, is clear: no model of the telecommunication network is needed, and the approach can be trained by using data coming from the field and/or by publicly available datasets. Moreover, due to the difficulty of providing precise models of the mobile network and all its components, AI-based tools may outperform model-based approaches and can be robust to model uncertainties.

Technical advancements and innovations progress beyond the state-of-the-art at the time of writing

NANCY will deploy a mechanism capable of conducting anomaly detection, self-healing, and self-recovery operations. The goal of this approach is to provide UEs and/or access networks with autonomous processes that can take decisions and prevent network faults on the device and network levels. In order to align the problem parameters to real-world conditions, we also consider the battery

of the UEs as an important optimization constraint. To this end, NANCY proposes the implementation of a Federated Learning (FL) framework that treats each device separately, according to its unique characteristics and preserves users' privacy. In such a framework, UEs are assigned a reinforcement learning (RL) model which is trained locally, without sharing any data or model parameters with the rest of the devices. During this process, the model learns the features of the corresponding device (battery capacity, energy consumption, computation resource, latency, etc.) and engages with proactive decision-making to prevent critical events such as high latency, low SNR, insufficient computational resources, and/or battery exhaustion. Through this scheme, each RL model will be locally trained and as a result, will adapt to the requirements of each device. In the sequel, the locally trained models will be dispatched to a centralized MEC node where the model aggregation process takes place. This will merge all the local models into one global model which, in turn, is sent back to the UE. This aggregation operation will assimilate the model weights to create a simpler, yet efficient, version of them. Further, it can impose new global constraints or rules to the optimization problem, if required. On the UE side, the newly acquired global model will be used for the local training process thus, triggering a new FL cycle. NANCY employs an FL approach mainly because there is no need for data exchange between devices and because device data is not sent to the network, but instead, a model trained on such data (which preserves users' privacy) is sent to the MEC. In this scenario, the only information exchange between the Edge and the UEs considers the model parameters which significantly increases the QoE of the end users.

Contributions towards the realization of the NANCY architecture

The Next-generation SDN-enabled MEC for autonomous anomaly detection, self-healing, and self-recovery module will be designed with respect to the following features: (**i**) high energy-efficiency; (**ii**) Data security and privacy; (**iii**) provisioning of edge intelligence to the UEs. Energy efficiency will be achieved by training the ML model to consider the remaining battery of the UEs. As a result, anomalies related to the low remaining UE battery will be identified and proper mitigation measures will be deployed. Moreover, the local model training process, which takes place on the UE side, will be designed to keep the computational complexity of the required operations to a bare minimum. In this way, UEs will save both energy and resources when training the ML model. Another strength of the proposed module is the fact that data security and user privacy are guaranteed due to the nature of the FL. User data stay confined within each UE and are not exchanged with any other node. Instead, UEs dispatch only the locally trained models to a sink node where the model aggregation takes place. This process ensures that data security issues are handled properly. Another characteristic of the envisioned module is the provisioning of Edge intelligence to the end-users. Contrary to existing approaches where the Edge is considered a monolithic entity that serves a number of UEs, NANCY employs a distributed scheme where both the UEs and the Edge participate in the training process. This approach gives the Edge the flexibility to adaptively serve multiple UEs and allows for the UEs to collectively take advantage of a large pool of computational resources for the training process. This type of intelligence would otherwise be inaccessible to the UEs due to the large computational requirements of most ML schemes. By leveraging the features analysed above, NANCY designs a component to solve the gaps in the Edge computing ecosystems and to further push the innovation limits.

Interconnections and dependencies with other results

The next-generation SDN-enabled MEC for autonomous anomaly detection, self-healing, and self-recovery components belongs to the Edge plane of the NANCY architecture. It receives input from the "novel AI virtualizer for underutilized computational & communication resource exploitation ([R8])" to assess the amount and type of computational resources available for the FL training. Also, it

communicates with the User plans where it distributes the global models to the UEs, during the FL process. This bi-direction communication is supported through the "Semantic & goal-oriented communication schemes for beyond Shannon performance" module ([R11]). For the purposes of ML model selection, the envisioned "Next-generation SDN-enabled MEC for autonomous anomaly detection" utilizes the ML models stored in the Novel self-evolving AI model repository ([R9]) and leverages the MLops to simplify and streamline the FL training.

3.3.2. [R14] A computational offloading mechanism with novel resource-aware/provision scaling mechanisms and novel battery as well as computational-capabilities aware offloading policies

Technical advancements and innovations state-of-the-art at the date of writing

Computation offloading in mobile edge computing promises to increase resource efficiency, improve performance, and minimize energy consumption. Towards this direction, TALON [179], a HORIZON EU project, implements an AI orchestrator that redistributes computational tasks between Edge computing nodes and devices in Industry 4.0 environments. The goal of this task scheduling is to minimize battery consumption and to balance the computational load between UEs and Edge nodes. 5GZORRO [135] which is a H2020 research project, employs a data-driven resource management mechanism that monitors the network in real-time and utilizes neural networks to derive decisions for the scheduling of tasks within the available devices. Network functions are considered critical in computation offloading schemes, as 5G-PICTURE demonstrates [180]. In this H2020 project, hardware resources are directly configured by network functions and thus, computing and storage resources are distributed on the fly in an efficient way. Offloading of software has also been realized through CPU reservations [181], which allow assigning different software components to different virtual CPUs with a guaranteed fraction of the overall computing capacity. This allows, for example, to more efficiently offload network functions that cannot saturate entire physical CPUs.

Generally, computation offloading techniques are considered key enablers of the 6G networks [182]. As a result, time-critical applications that require ultra-low delays or high QoE, seem to enjoy most of the benefits of this approach. Examples of such applications usually include Vehicular Edge Computing [183], IoT paradigms [184] and ML [185]. Aside from the application side, several works focus on the algorithmic side of computational offloading. In this area, previous research considers Nash equilibrium modeling [186], Markov decision processes [187], and machine learning [188] to optimize the offloading efficiency.

Purpose of the component development in view of gaps defined in 2.3

This result focuses on providing the NANCY system with the necessary capabilities to be able to dynamically cope with an increased demand for system features, such as computational capacity, security, latency, or bandwidth. NANCY is a project characterized by the presence of multi-operator scenarios, where the workload forecasts of the nodes are exposed to peaks of high demand. This is why having an offloading capacity within distributed nodes is essential to guarantee the stability and performance of the system. The selection of the optimal node, as well as the monitoring of node resources and load forecasts, are fundamental tasks linked to this result. This outcome is a central piece of the system since the offloading capability will be used by a vast majority of the services to be developed (intra and inter-operator), e.g., the orchestration capabilities enable automatic lifecycle management of the services deployed on the MEC nodes, including resource scaling or transparent mobility of the services. Furthermore, this result has a strong dependency on the AI engines to be developed, since in addition to decision-making, the engines can be deployed in the MEC to support a wide range of processes.

Technical advancements and innovations progress beyond the state-of-the-art at the time of writing

NANCY will deploy machine learning functionalities to support the dynamic computation offloading operations of the platform. The consortium will devise a context-aware offloading mechanism that will pro-actively transfer computations to the edge of the network. The goal of this process is to maximise resource utilization among the available devices while also considering the capabilities of MEC nodes. In contrast with existing approaches, NANCY will formulate a multi-variate optimization problem that considers several parameters of the network such as user preferences, radio, and backhaul connection quality, and UE capabilities. Meeting the optimization objective of network functions also depends on the performance of the computational threads implementing them; therefore, their performance will be carefully considered. To solve the optimization problem, the [R14] will employ a combination of machine learning, optimization, and feature extraction techniques. Initially, the network parameters will be concatenated into a unified data space. In the sequel, the feature extraction technique will expand the parameter space and will identify the important data features that are estimated to significantly affect the model performance. Using such features, a machine learning model will be implemented and will search for an optimal solution to the problem. This optimum represents the number of computations that should be offloaded to the corresponding nodes. The optimization procedure will also leverage the allocation of workloads to virtual CPUs, considering resource reservation mechanisms capable of providing a virtual CPU abstraction while guaranteeing CPU quota and maximum service delay and allowing multiple applications to execute on the same groups of CPUs, thus avoiding the natural system under-utilization that occurs when having a one-to-one mapping between virtual and physical CPUs. By leveraging the feature extraction mechanism prior to the machine learning model, NANCY will be able to consider a significant amount of network parameters for the decision-making process; without increasing the problem complexity by a large factor.

Contributions towards the realization of the NANCY architecture

The task offloading mechanism is one of the central functionalities of the project; aided by the artificial intelligence engines, it seeks to maintain the optimal use of resources, thus contributing to maintaining responsible energy consumption while ensuring adequate levels of Quality of Service and Experience. Task offloading mechanisms may be required intra or inter-operator, deploying network and processing functions in the best possible node to offer advanced services with such expected and maintained quality in terms of security, latency, bandwidth, connectivity, reliability and computation, among others. Task offloading will also be supported by user-centric caching mechanisms (R15), allowing the mobility of virtualized services, deployed in different nodes, in a transparent way for the end consumer, a capability especially important in mobile scenarios or for handing over resources to other nodes. In addition, Task offloading involves a series of operations that should be tracked by the use of a blockchain., The mechanisms developed for this purpose will be used by operators to request services from others. Three types of task offloading policies will be developed to support NANCY architecture, namely, local offloading, full offloading, and partial offloading. Given the resource-aware mechanisms and the type of devices, the optimal policy will be selected. In particular, envisioned mechanisms will play a principal role in D2D capabilities, where connectivity, caching, and computation are three features that will be offloaded to provide availability of E2E services. One of the critical objectives of NANCY is to ensure reduced latency and high computational capabilities; to this aim, Task offloading mechanisms at the edge of the network are the technological propeller to provide this ambition.

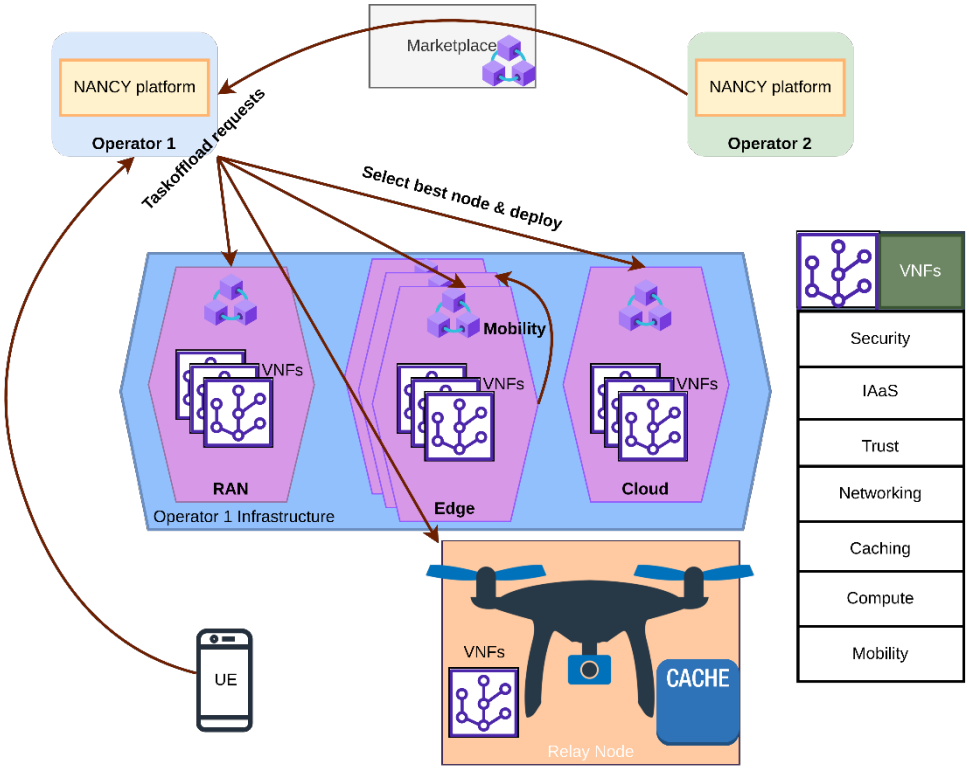

Figure 13 Task offloading scheme

Interconnections and dependencies with other results

In the following a description of the relation of R14 with other results is given:

- **[R1]** Task offloading is one of the central pieces related to the architectural realization of NANCY. Depending on the demanded task, the Cloud or the MEC may be adequate places to perform the required operations to achieve desired features.
- **[R2]** Task offloading in D2D scenarios is crucial for the correct performance of the connection, caching functionalities and mobility, which may be triggered by task offloading requests.
- **[R7]** A proper orchestration scheme is a prominent enabler to achieve the automation of task offloading.
- **[R8]** Deployed and offloaded VNFs are resource monitored to grant resource usage optimization.
- **[R10]** The models also include the selection of the nodes for task offloading, their scalability and the allocation of resources.
- **[R13]** SDN is closely related to the autonomous deployment of the offloaded tasks, this task considers security to networking features as well as computing activities.
- **[R15]** Certain tasks could have increased performance if user-centric data is brought close to the end-user, to this aim, task offloading will take advantage of cached content to deliver faster and better services.

3.3.3. [R15] User-centric caching mechanisms

Technical advancements and innovations state-of-the-art at the date of writing

The advent of Edge Computing (EC) along with the deployment of 5G networks has led to new communication models that require higher data rates, low latencies, and efficient power utilization. Standard network topologies cannot achieve such results, since the generated traffic puts enormous pressure on the network backhaul, and thus, lower data rates are expected. To address this issue, caching mechanisms are often employed which pro-actively transfer content (or services) to the Edge of the network. Within this context, user requests are served from the Edge, instead of the cloud/BS. As a result, a question arises on what (content) to cache and when to cache at the Edge. Previous works in the area suggest that pro-active content caching should also consider the energy consumption of the UEs [189] [190], while other works focus on efficiency in terms of cache hit rate [191] [192]. Regardless of the objective function (either energy or latency) of the researchers, there is a growing trend of using ML techniques to proactively deploy services and offload data to the Edge [193]. More specifically, RL methods are often employed due to their capacity to provide robust results in cooperative multi-agent environments [194]. D-RL techniques also show promise, in terms of cache hit rate, especially in Device-to-device networks (D2D) [195]. From the D-RL domain, Deep Q-Learning (D-QL) algorithms [196] seem to achieve state-of-the-art performance since they can be used in dynamic network ecosystems with various content popularity, where different caching polices co-exist.

On the EU-funded research side, 5G-Xcast [197] is an H2020 project that focuses on media delivery and media caching techniques. 5G-Xcast develops novel multicast modes that aim to increase the QoE of the end-users, by serving them media cached at the edge of the network. CacheMire [198] is an ERC project that employs collaborative QL approaches to optimise data caching in D2D networks. In CacheMire data is not necessarily cached at the Edge of the network; instead, it is stored on several user devices in a distributed way. Then, when a consumer requests data from the network, the envisioned QL technique is put into play which designates the device which will provide the data to the request. PriMO-5G [199] develops a data caching strategy for consecutive user demands in categorized contents. To accomplish this, researchers utilize schemes where they maximize the success of probabilities for content delivery of all users.

Purpose of the component development in view of gaps defined in 2.3

This result is closely linked to the achievement of near-zero latency connectivity. Future 6G services will have to be tailored to service and customer requirements. Among these requirements, latency is one of the most constraining components, requiring accurate predictive models that load the heaviest data and processes in advance, in order to be ready in advance of their use. This will not only reduce service consumption time but also energy consumption and compute load distribution will be improved by these caching methods. In particular, user-centric caching techniques aim to focus on preparing the infrastructure, processes, and dataset necessary to improve the user's service experience. User mobility is one of the main characteristics to be taken into account in the design of these mechanisms, facilitating the transfer of the required services to neighbouring nodes, e.g., at the edge, potentially visited by the end user.

Technical advancements and innovations progress beyond the state-of-the-art at the time of writing

NANCY envisions a social-aware user-centric caching mechanism, where the profiling of each user will provide the system with adequate data to efficiently predict and pro-actively offload data to the Edge of the network. To accomplish this, the content request history of each user will be monitored and will be fed to an RL model. This model will perform inference over the user requests (within a time window) and, through a reward function, it will make predictions on future data requests. In the sequel, NANCY will offload the selected data to the Edge, or to the nearest devices, so that to minimize E2E

transmission latency and to alleviate the backhaul pressure of the network. Different AI models will be explored which will be tailored to the current network conditions. For example, when the network's backhaul, traffic is measured high, the AI models will adjust to compensate and thus, the data offloading process will be performed within stricter constraints. On the contrary, in situations where network traffic is low, the RL will employ looser criteria for the caching process.

Contributions towards the realization of the NANCY architecture

Supported by AI engines that select the best node to bring the content and the content to be cached, user-centric caching mechanisms are a crucial element to achieving almost-zero latency connectivity but also as enablers of D2D communications. Envisioned mechanisms will provide: i) reduction of the data retrieval time, storing frequently accessed content at the closest node of the edge, and preparing (e.g., instantiating) services and procedures that could be also reallocated in closer proximity; ii) lowering network congestion, as the mechanisms reduce the average time that the data must travel around the network to arrive to the final destination, avoiding centralised stations such as cloud centers, thus leveraging distributed architectures; iii) enhancing D2D communications, preventing the access to central cloud to deliver services, reducing latency and improving the quality of service and experience; iv) optimized content localization, leveraging the rest of the highlighted points, based on user and network patterns, allocate the necessary resources and content into the predicted nodes, recurring to the mobility of the cached content along with the mobility of the user; v) real-time data-access, user-centric caching mechanisms are conditioned by network changes, they are adapted and can respond to real-time events, ensuring that the content is ready to be consumed at any moment, which improves the reliability and responsiveness of the system, especially in D2D scenarios; and vi) overall QoS improvement, well designed caching mechanisms that can select high-priority content are proved to enhance the overall QoS performance, maintaining low-latency and uninterrupted performance.

Interconnections and dependencies with other results

- **[R2]** User-centric cache mechanisms are essential blocks for D2D, given that the connection should be retained as close as possible to the source to reduce latency and provide high-quality services.
- **[R9]** The cached content is evaluated to check for improvement in various areas of the system in order to detect which content is of high priority for certain services. Thus, evolving the AI learning processes.
- **[R14]** Task offloading mechanisms are analysed rigorously to identify which content is categorized as high priority to the requested task and, thus could be optimized by allocating it at a closer edge node.

4. NANCY Overall Architecture

In the NANCY project, it is anticipated that the implementation of an AI-optimized B-RAN architecture will facilitate the development of secure and private B5G networks. This advancement will go beyond localized enhancements limited to the radio access level, and instead, bring about a comprehensive transformation in which the network will seamlessly integrate data processing, storage, and transport connectivity functions. In addition to the imperative need for security and privacy measures, there are other crucial attributes that are essential for the successful implementation of B5G networks. These attributes are indispensable in order to effectively handle transformative applications. They encompass flexibility, adaptability to dynamic topologies, the ability to accommodate extreme wireless node densification, low-latency communication, ultra-high reliability, energy efficiency, and high-data rates. The NANCY framework incorporates the anticipated applications and qualities of B5G technology. As illustrated in Figure 14, NANCY places particular emphasis on three meticulously chosen representative usage scenarios, namely (a) fronthaul of fixed network topology, (b) advanced coverage expansion, and (c) advanced connectivity of mobile nodes.

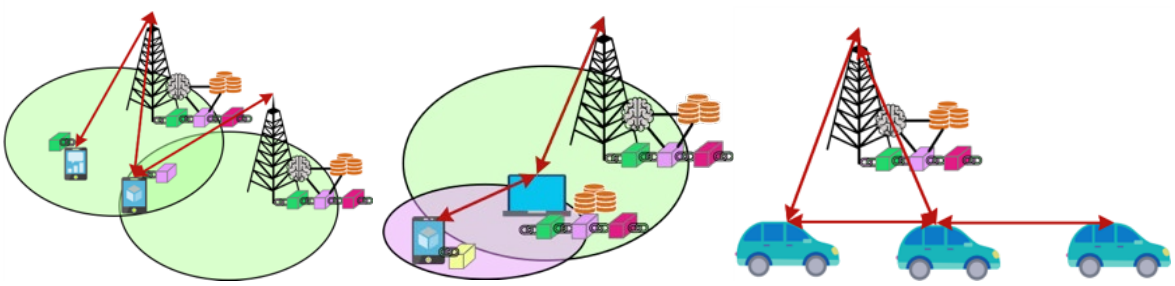

Figure 14 NANCY usage scenarios

4.1. NANCY platform architecture

NANCY aims to create a flexible, scalable, and energy-efficient platform for B5G networks, ensuring high security and privacy. This requires commercializing underutilized resources, adapting novel network technologies, and rethinking conventional network design principles. It focuses on AI-aided Blockchain wireless radio access B5G networks, which require a flexible, scalable, and powerful ML-based orchestration framework, novel blockchain and attack models, a revolutionary network information theory approach, and cutting-edge technology components. The approach of NANCY is built upon three complementary pillars, i.e.,

- Pillar I: Distributed and self-evolving B-RAN for dynamic scalability, high-security, and privacy in a heterogeneous environment;
- Pillar II: Pareto-optimal AI-based wireless RAN orchestration for energy efficiency and trustworthiness; and
- Pillar III: Distributed MEC for almost-zero latency and high-computational capabilities at the edge.

The NANCY platform is developed in line with the B5G network evolution aiming to exploit underutilized resources and design innovative concepts for the proliferation of new DoFs. In this context, the NANCY project investigates, designs, develops, and validates the NANCY B-RAN

architecture, focusing on AI-based architecture built on SDN and MEC principles. The goal is to meet the requirements of the B5G era by proposing an AI-based architecture, accurately modeling and assessing its performance by means of respective Monte Carlo simulation theory, as well as in real-world conditions, inventing new technologies and system concepts, and validating and optimizing them through five (5) testbeds/demonstrators.

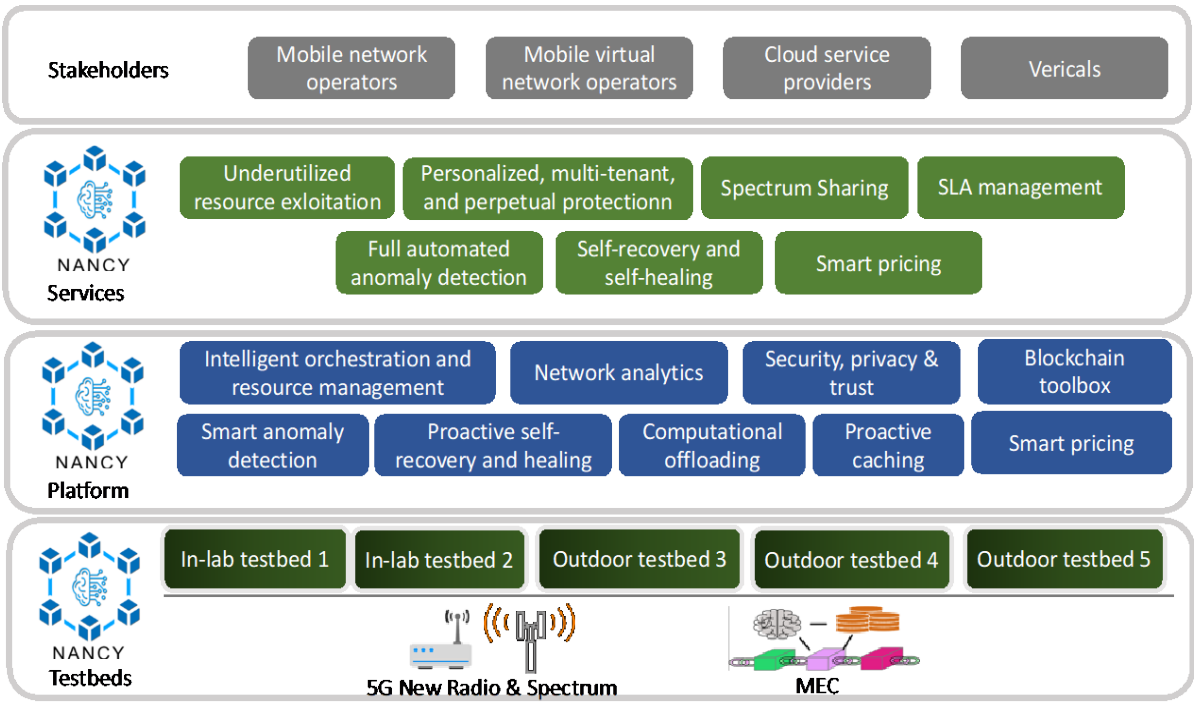

Figure 15 NANCY platform architecture

As illustrated in Figure 15, the NANCY components include enabling common network functionalities, blockchain and cell-free radio access mechanisms, AI-based resources, and network orchestration, distributed and decentralized blockchain approaches supported by MEC, and proactive self-recovery and self-healing mechanisms. In more detail, Fig. 15 presents the ecosystem of NANCY as a whole. The layer of interest from the architecture point of view is the NANCY platform, which is the technology enabler of the innovative services that NANCY is expected to support. The NANCY platform contains the following modules: i) intelligent orchestration and resource management, ii) network analytics engine, iii) security, privacy and trust mechanisms, iv) blockchain toolbox, v) smart anomaly detection techniques, vi) proactive self-recovery and healing strategies, vii) computational offloading mechanisms, viii) proactive caching approaches, and ix) smart pricing engine. These modules will be integrated into the O-RAN architecture in order to transform B5G RANs into intelligent platforms, opening new service models to telecom/ISP and individual providers. It will also identify critical technology gaps and invent, optimize, demonstrate, and assess key enablers and future directions for the B5G RAN.

4.2. NANCY high-level network architecture

B5G/6G networks are expected to deliver a wide range of services across various vertical sectors, requiring rapid resource allocation and network orchestration. They are highly distributed, requiring technologies like cloud-edge computing, SDN, and NFV, increasing their complexity. To address these challenges, incorporating a decentralized network like Blockchain in distributed networks can provide seamless services to end-users with transparency, security, and reliability.

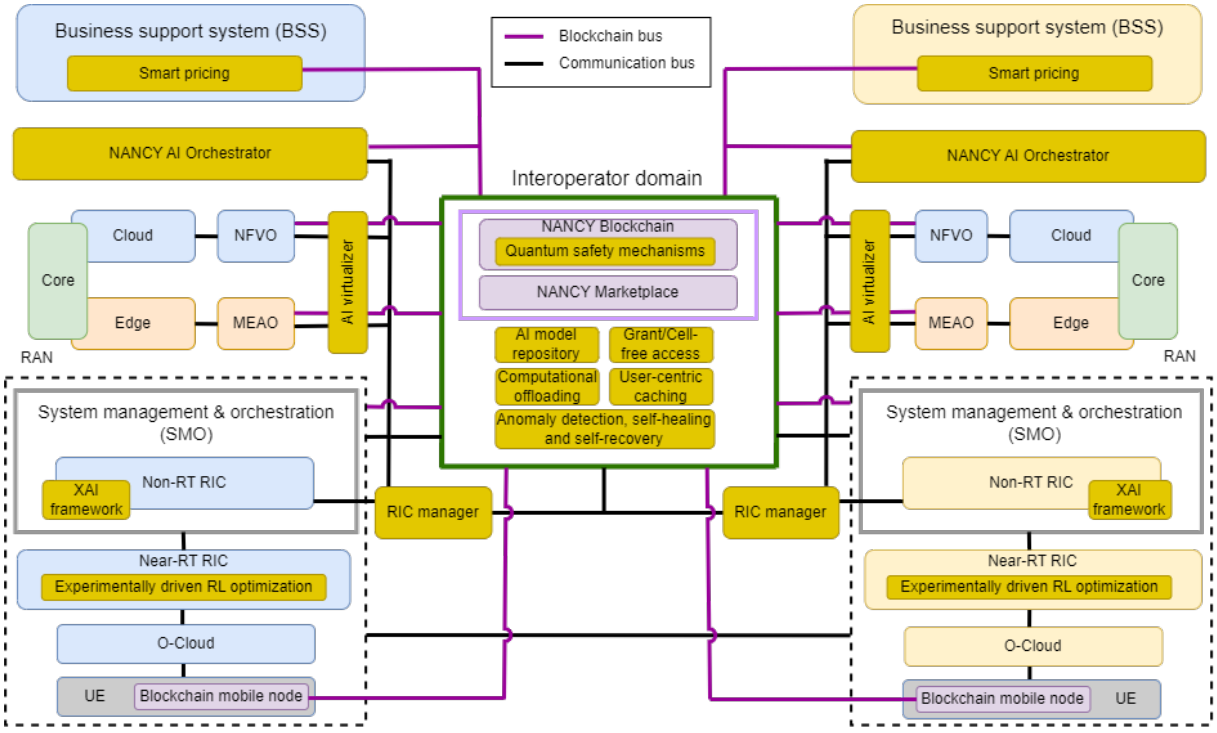

Figure 16 NANCY high-level network architecture

The NANCY high-level network architecture is depicted in Figure 16. The interconnections between the NANCY architecture and the various components of the core, edge, cloud, and RAN are presented. The core is deployed in the edge-to-cloud continuum, with the edge being governed by the mobile edge application orchestrator (MEAO) and the cloud by the network functions virtualization orchestrator (NFVO). Moreover, the RAN part of the architecture is denoted with the black dashed line and is based on the O-RAN architecture. The NANCY components are presented with yellow and purple denoting the network and Blockchain-oriented parts, respectively.

The key enabling technology of NANCY is the blockchain, which is accessible to the various end users of the NANCY platform through the marketplace that resides in the interoperator domain. The interoperator domain is the central plain of the NANCY architecture and houses the components of NANCY that are deployed and shared between different providers. Such components include the AI model repository, computation offloading, user-centric caching, as well as quantum safety, grant/cell-free access, anomaly detection, self-healing, and self-recovery mechanisms.

To ensure uninterrupted connectivity and seamless movement of UEs, multiple mobile relay nodes are deployed to establish multi-hop networks. This ensures data hops, network reliability, and coverage in challenging environments. The network infrastructure provides radio interfaces for mobile devices and vehicles. An exhaustive control and pricing system, incorporating blockchain systems and smart pricing

policies, facilitates interoperability, security, trust, and efficient resource management. This collaboration enhances user experience and collaboration between operators.

NANCY aims to address challenges in traditional O-RAN architectures by integrating cell-free access, multi-hop networking, and relay nodes. This approach improves resource allocation, synchronization, and scalability, allowing operators to adjust the number of RAN nodes as needed. Multi-hop networks efficiently redirect traffic to suitable RAN nodes, avoiding congestion. Operators collaborate to enable seamless connectivity between nodes, and blockchain systems will be used for secure transactions and authentication. This integration will also enable intelligent policies for resource sharing.

The NANCY system is designed to handle increased demands for features like computational capacity, security, latency, and bandwidth. The project is multi-operator, with high demand peaks affecting node workload forecasts. User-centric caching mechanisms are crucial for NANCY, ensuring connections are kept close to the source to reduce latency and provide high-quality services. The cached content is evaluated to identify high-priority content for services, evolving AI learning processes. Intelligent orchestration, resource management, and task offloading mechanisms are also analyzed to optimize high-priority content allocation to requested tasks, potentially at a closer edge node. Offloading capacity within distributed nodes is crucial for system stability and performance. The offloading capability will be used for most services, including orchestration capabilities for automatic lifecycle management and resource scaling. The result heavily relies on AI engines for decision-making and support for various processes in the MEC. The offloading capability is central to the system's success.

Each individual operator is connected to the interoperator domain in order to take advantage of the aforementioned functionalities offered by NANCY. In this direction, some NANCY components must be deployed on the premises of each operator and therefore will have multiple instances throughout the overall NANCY platform. Such components include the NANCY AI orchestrator, the AI virtualizer, the RIC manager, the experimentally driven RL optimization mechanisms, the XAI framework, and the smart pricing policies. These modules are necessary for supporting the necessary functionalities of the NANCY platform. Specifically, the AI orchestrator of each operator communicates not only with the components of the interoperator domain, like the marketplace, computational offloading, caching, and other, but also with the RIC manager and the AI virtualizer, in order to achieve the optimal instantiation and management of the resources of its operator. Furthermore, the smart pricing module is situated in the business support system (BSS) and provides the marketplace with the necessary policies for creating and updating the smart contracts before being deployed in the blockchain. Finally, the XAI framework communicates with the non-RT RIC in order to provide explainable insights into the various aspects of the network operations.

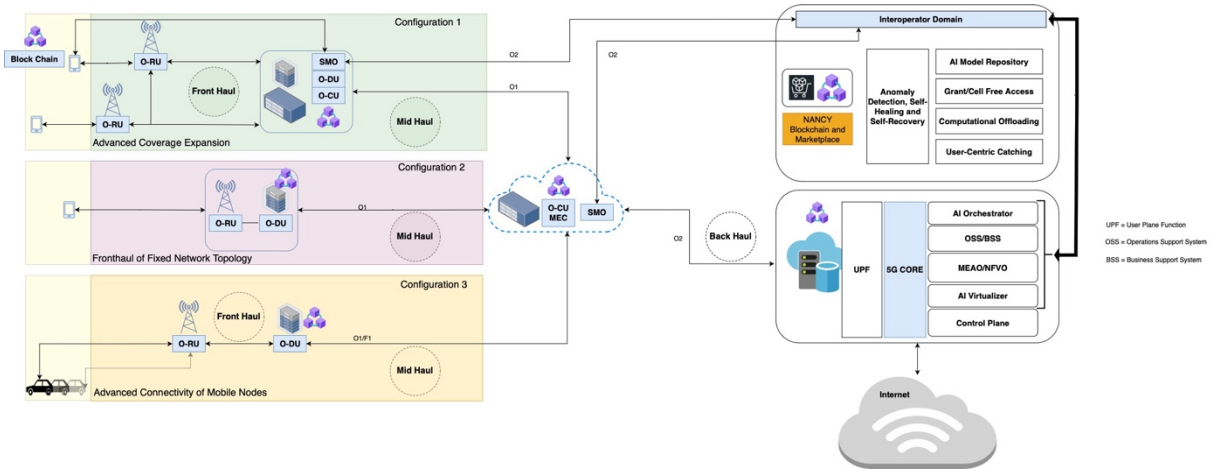

Figure 17 Flexibility of NANCY architecture using different configurations.

Next-generation wireless networks need a major paradigm shift from the conventional network design. NANCY project incorporates B-RAN architecture and multi-hop systems to adapt to the fast-changing communication environment while providing ultimate security in the B5G era. The decentralized design of B-RAN allows O-RU, O-DU, and O-CU to be distributed in the networks, and the benefits of this approach are multi-fold. For example, the scalability is expanded due to the nature of B-RAN, where B-RAN components can be deployed freely in the networks upon the operator's requirements, allowing the operators to provide extensive coverage and overall good signal strength for users. More details are discussed in Section 3.1. Moreover, multi-hop networks can be established to avoid data congestion and ensure seamless communication for moving nodes with the help of the MRAT-NCPs and NANCY authentication, which enables the signals to be relayed from one node to another as long as the mobile equipment is granted by the NANCY Blockchain mechanism. The discussion can be found in Section 3.1.2. The aforementioned benefits are illustrated in Figure 14 "NANCY usage scenarios." while Figure 17 shows the flexibility of the NANCY architecture with different configurations. Furthermore, NANCY deploys distributed AI-powered elements among different domains in the network. These components are lightweight and capable of self-optimization/healing, which can bring a high degree of automation and short deployment time to the network plan. More particularly, AI-based orchestrators can harness the power of machine learning techniques to develop solutions for network slicing in highly dynamic environments swiftly through AI virtualization. These so-called AI virtualizers can enhance the flexibility of resource managers by identifying the computational resources and making offloading decisions. Additionally, the orchestrators are able to deliver flexible management of different layers and perform resource sharing between operators through the NANCY blockchain and marketplace in the interoperators domain; the detailed architecture is shown in Section 4.3.

4.3. NANCY B-RAN architecture

The B-RAN architecture focuses on the RAN part of the mobile network, which connects devices to the core network infrastructure, and the Blockchain, a distributed ledger technology that ensures secure and transparent record-keeping through a decentralized network of nodes. Smart Contracts, self-executing contracts stored on the Blockchain, facilitate automated and secure interactions between network participants. The B-RAN architecture aims to enhance security and privacy in the management and operation of mobile networks, with Blockchain's decentralized and tamper-resistant nature providing protection against unauthorized access and data manipulation.

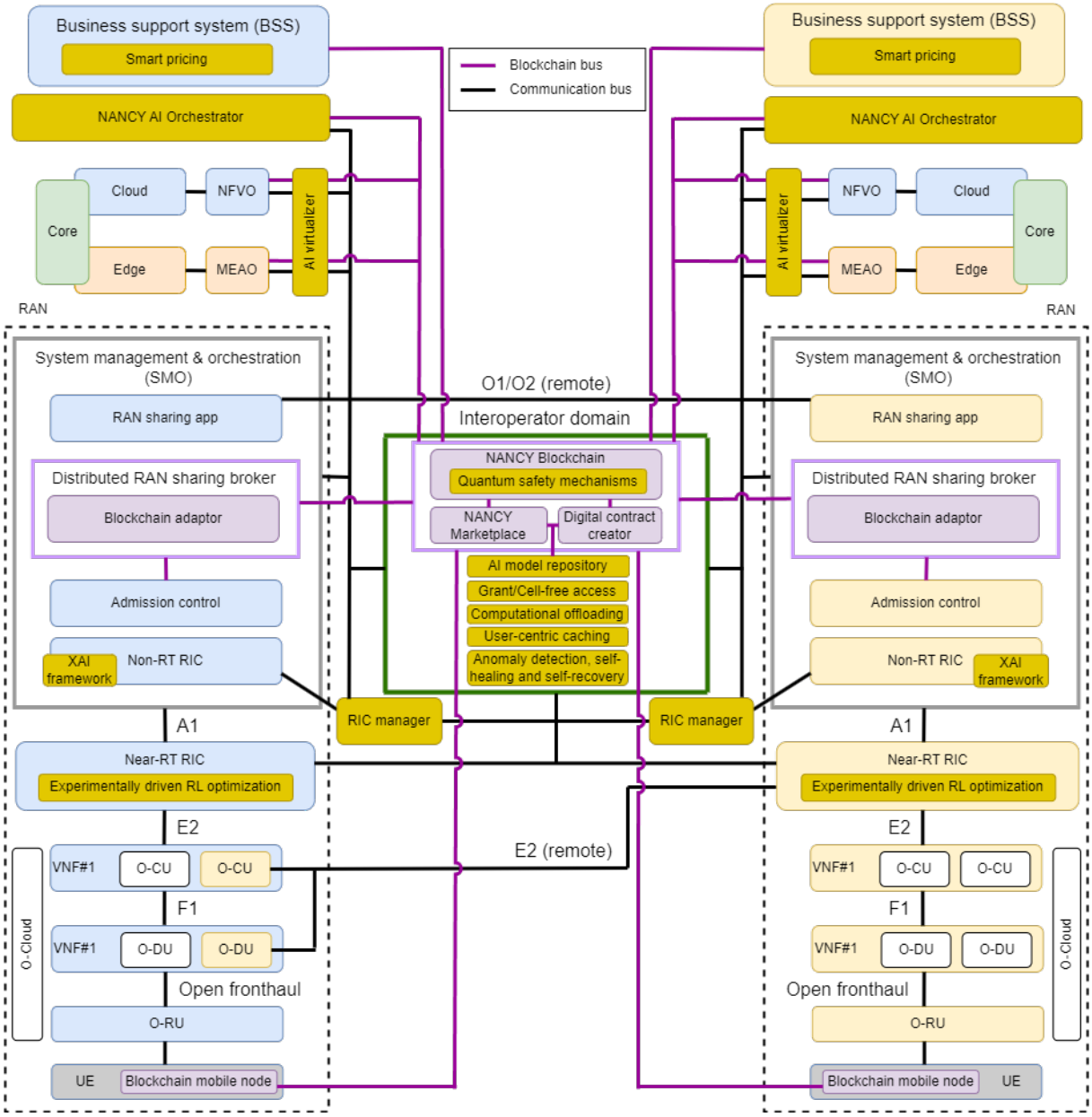

Figure 18 NANCY B-RAN architecture

4.3.1. Blockchain

The NANCY Blockchain is a fundamental element of the B-RAN architecture illustrated in Figure 18. It aims to enhance user privacy and security by combining anonymous credentials, such as SSI, with open-source implementations like Hyperledger Fabric. The proposed SSI infrastructure should have a data registry for cryptographic material, and the NANCY blockchain can serve as the data registry. The blockchain offers flexibility in implementing various solution models, including account models, UTXO models, structured and unstructured data, and Turing complete smart contracts. It also supports data privacy through channels or private data collections. NANCY's blockchain is designed for continuous operations, including rolling upgrades and asymmetric version support. Moreover, it provides governance and versioning of smart contracts and a flexible endorsement model for consensus across organizations. Another innovation offered by the NANCY blockchain is the protection of smart contracts against attacks, such as reentrancy attacks. Mitigation measures include static analysis, language-based security, and runtime verification. NANCY will also analyze and propose lightweight primitives for permissioned Blockchain. Users will be provided with a wallet that enables interaction with the blockchain, featuring post-quantum security.

4.3.2. PQC

All NANCY devices, either providers or consumers of resources, are authenticated and registered on the NANCY blockchain, allowing access to the ledger. End-users authenticate onto the NANCY blockchain using the PQC means. The public part of PQC credentials is stored on the NANCY blockchain. The PQC solution that is provided by partner TDIS, is likely the main dependency as it will enhance security for devices connected to the NANCY Blockchain, enabling the Blockchain to check transaction signatures.

TDIS introduces PQC digital signature in the NANCY architecture to ensure blockchain security resilience. Current blockchain technology uses Public-key cryptography and hash functions, but SHA-256 is considered quantum-safe. To address this, Public Key cryptography must be replaced with a quantum-resistant scheme. Public-key cryptography establishes a distributed consensus of trust, but wallets at endpoints can be hackable. To address this issue, quantum-safe crypto wallets secured by PQC digital signature are proposed.

TDIS develops the Crystals Dilithium SHAKE with Security Level 3 for the PQC Digital Signature component within the NANCY project. The innovation will focus on maintaining requirement constraints while implementing a stronger PQC algorithm, ensuring a cost-effective implementation on small CPU devices, implementing high-security cryptographic algorithms with countermeasures against advanced attacks, and achieving acceptable performance compared to classical cryptography. Additionally, TDIS aims to introduce a novel mechanism for Crypto Agility.

4.3.3. Smart pricing

The Marketplace is a set of providers and consumers that exchange goods and services through smart contracts, maintaining authenticity, integrity, and privacy. The smart pricing module periodically updates smart contract data, connecting it to the Blockchain during operation. NANCY's smart pricing policies should incorporate security and privacy mechanisms, be scalable for large mobile users, and consider game theoretic pricing schemes to balance strategy and user choices. To achieve this, NANCY will integrate Blockchain into the RAN, enabling smart policies and AI techniques to provide monetary user incentives and regulate resource sharing. Offloading incentives, such as discounted access fees and token rewards, will be considered. Reputation-based incentives for users who consistently

contribute resources will also be considered, with higher reputation scores granting access to premium services, additional rewards, and tiered pricing based on resource contributions.

4.3.4. Grant/cell-free cooperative access

NANCY's architecture uses cell-free cooperative access mechanisms and multi-hop networks to improve mobile networks and B5G. These networks facilitate data transmission through multiple relay nodes, improving coverage, capacity, and efficiency. They also facilitate efficient traffic redirection to prevent congestion. NANCY addresses connectivity gaps and congestion by deploying relay nodes. It also enables UE to have unrestricted mobility, allowing them to connect to operators they don't belong to, ensuring uninterrupted and seamless connectivity. These mechanisms interact with the B-RAN architecture through lightweight consensus mechanisms, decentralized blockchain components, and smart pricing policies. As a result, data integrity and privacy are ensured through blockchain systems, recording transactions and agreements between operators.

Enhancing the O-RAN architecture requires focusing on dynamic confirmation control and quantum-resistant encryption methods. Dynamic confirmation control can improve network security by adjusting the trade-off between latency and security, while quantum-resistant encryption and trustless consensus procedures can enhance security. Real-time QoS monitoring, blockchain-based reputation systems for Access Points, smart contracts for dynamic QoS control, and decentralized auditors for public performance can also improve quality assurance. These developments will define B-RAN, ensuring greater security, robust service quality, and adaptability to advanced technologies.

4.3.5. AI-based orchestration

NANCY is leveraging AI-based orchestration capabilities to autonomously manage radio, network, and computation resources in network slices, aiming to achieve E2E service delivery. The system aims to enable inter-operator resource orchestration, enabling resource sharing between operators independently of network segments and layers. Blockchain plays a crucial role in radio resource orchestration, providing trust mechanisms for various purposes. AI engines control and enhance orchestration procedures, such as optimizing radio resource allocation and placement. Reinforcement learning allows for an AI solution to adapt to dynamic changes and use advanced techniques like transfer learning. This allows for flexible and elastic network slicing, allowing the solution to quickly adapt to the deployment environment. For example, an AI-based slicer evaluates specific requirements for an application, such as bandwidth, latency, and processing power, and selects the required resources to instantiate the network slice. Over time, the agent learns to optimize network performance.

4.3.6. Self-evolving AI model repository

In the same direction as the orchestrator, the introduction of a self-evolving AI model repository will significantly improve AI-native RAN systems by reducing inference time and optimizing response times. NANCY aims to develop automated approaches to feature selection, model search, and selection based on ML/AIOps principles and tools. The project uses the NANCY common data model, which allows features to be selected, combined, and engineered using classical feature interaction and enrichment approaches or representation learning techniques. The pipeline also enables uploading and training contributed models on-demand or scheduled using neural architecture search and hyperparameter optimization techniques. This shift from monolithic models to dynamic model selection represents a new era in machine learning. The self-evolving AI model repository champions specialization, tailoring each model to excel in specific tasks or environments. This approach provides a more streamlined and

efficient path for AI model deployment and administration. The NANCY platform aims to bolster its adaptability and versatility by enabling fluid retraining and updating models with diverse datasets and facilitating feature extraction from substantial data volumes.

4.3.7. Experimentally driven RL optimization

The experimentally driven RL optimization of B-RAN aims to improve intelligent components in O-RAN networks. This component will focus on fast, real-time decision-making, achieving ultra-low latency delays. The model will consider noisy or incomplete data and use feature extraction techniques to alleviate penalties. The module will expand input data dimensionality through multiple layers of feature extraction operations, feeding extracted features to the ML model. Data gaps or inconsistencies will be considered low-quality features and play a trivial role in decision-making. To increase model adaptability to complex environments, NANCY will consider early-exit-inference mechanisms, which enable the AI model to stop the inference process according to network conditions or functional requirements. This will be particularly useful in lower resource variability environments, where the model can generate predictions earlier.

NANCY is developing a re-trainable optimization framework for resource allocation, using reinforcement learning (RL) techniques. The framework will use an agent-based modeling approach to select the optimal number of resources in real-time. The agent engages in a decision-making process and is rewarded for each decision. The dynamic reward function, which changes according to network conditions, will enable faster convergence to the optimal decision point. This approach merges the learning rate of the neural network with the RL's reward function, achieving faster convergence rates and making it efficient for real-time systems requiring quick decision-making.

4.3.8. AI virtualizer

The experimentally driven reinforcement learning optimization of the B-RAN module is linked to the AI orchestrator, the novel AI virtualiser for underutilized computational and communication resource exploitation, and the novel self-evolving AI model repository. The AI orchestrator provides network organization details. The self-evolving model repository updates its internal state to keep its internal state in sync with network conditions. The repository also engages in model retraining actions to improve stored models' performance. Finally, the experimentally driven reinforcement learning optimization of B-RAN constantly requests optimized network versions from the AI virtualizer, which interacts with other NANCY modules, managing ML-OP and lifecycle with the help of the novel self-evolving AI model repository.

O-RAN uses machine learning to enhance network adaptability, efficiency, and responsiveness to user needs. However, it lacks a fully operational intelligent orchestrator integrated with ML technologies. The proposed AI virtualizer in NANCY aims to improve the orchestrator's functionality by incorporating ML interactions, optimizing resource utilization, and enhancing resource manager adaptability. Specifically, it identifies necessary computational resources and makes offloading decisions, allowing for the exploitation of unutilized resources. The virtualizer aims to mitigate inter-slice conflict and minimize underutilization by intelligently using spare resources from concurrent slices without exchanging monitoring data. Based on a multi-agent deep reinforcement learning (MA-DRL) communication framework, each slice has one agent responsible for resource orchestration via CPU scaling. DRL agents learn and discover signaling policies cooperatively, guided by a reward function that penalizes conflicts and underutilization while minimizing latency.

4.3.9. Task offloading

Towards the same direction, task offloading is a crucial aspect of NANCY architecture, enabling the correct performance of connections, caching functionalities, and mobility in D2D scenarios. A proper orchestration scheme is essential for automating task offloading. Deployed and offloaded VNFs are monitored for resource usage optimization. Models include selecting nodes for task offloading, scaling, and resource allocation. SDN is closely related to autonomous deployment of offloaded tasks, considering security, networking features, and computing activities. Task offloading can increase performance by bringing user-centric data closer to the end-user, utilizing cached content to deliver faster and better services.

Task offloading mechanisms are a key component of the project that contribute to responsible energy consumption and ensure quality of service and experience. These mechanisms can be used intra or inter-operator, deploying network and processing functions in the best possible node to offer advanced services with expected quality in security, latency, bandwidth, connectivity, reliability, and computation. User-centric caching mechanisms support the task offloading, allowing transparent mobility of virtualized services in different nodes. Task offloading operations are tracked using a blockchain, and three types of policies will be developed to support NANCY architecture: local offloading, full offloading, and partial offloading. These mechanisms will play a crucial role in D2D capabilities, ensuring reduced latency and high computational capabilities.

4.3.10. Social-aware caching

Alongside task offloading, NANCY's social-aware caching will utilize user profiling to efficiently predict and offload data to the network's Edge. The system will monitor user content request history and feed it to a reward function, which will perform inference to predict future data requests. The data will then be offloaded to the Edge or nearest devices to minimize E2E transmission latency and alleviate network backhaul pressure. Different AI models will be explored to adjust to network conditions, such as high backhaul traffic or low network traffic, adjusting the data offloading process accordingly.

These mechanisms will also focus on D2D communications to prevent access to central cloud, reducing latency and improving service quality. Optimized content localization is achieved by allocating resources and content into predicted nodes based on user and network patterns. These mechanisms are conditioned by network changes and can respond to near real-time events, ensuring content is ready for consumption at any moment. Overall, well-designed caching mechanisms that select high-priority content enhance overall QoS performance, maintaining low-latency and uninterrupted performance.

4.3.11. Anomaly detection, self-healing, and self-recovery

The next-generation SDN-enabled MEC for autonomous anomaly detection, self-healing, and self-recovery module will focus on high energy efficiency, data security and privacy, and provisioning of edge intelligence to the UEs. The ML models will be trained considering the remaining battery of UEs, identifying anomalies with low remaining battery and deploying mitigation measures. The local model training process will minimize computational complexity, saving energy and resources. Data security and user privacy are guaranteed due to the FL's nature, with user data confined within each UE and dispatched only to a sink node for model aggregation. The module also provides edge intelligence to end-users through a distributed scheme, allowing the edge to adaptively serve multiple UEs and collectively utilize a large pool of computational resources. This innovative approach aims to solve gaps in edge computing ecosystems and push innovation limits. To achieve this, this module will take

advantage of the novel AI virtualiser to assess computational resources and will capitalize on the trained models of the self-evolving AI model repository.

4.3.12. XAI framework

Finally, the XAI framework will provide non-real-time explanations and be accessible to NANCY users with the required authorization. It will communicate with the non-RT RIC and it is closely linked to the self-evolving AI model repository aiming to support the interpretation and explainability of decisions and predictions made by novel AI models. The 4-stage XAI engine uses SHAP and LIME technologies to explain local model behavior and provide global insights into the importance of specific features. The SHAP algorithm extracts SHAP values, which are used for descriptive elements like feature dependencies, explanations, and models' summarization. The LIME technique then produces local explanations, feature importance, and model evaluation. The outputs from the second and third stages are fed into the final stage for diagnosis and decision-making tasks.

4.4. NANCY orchestration architecture

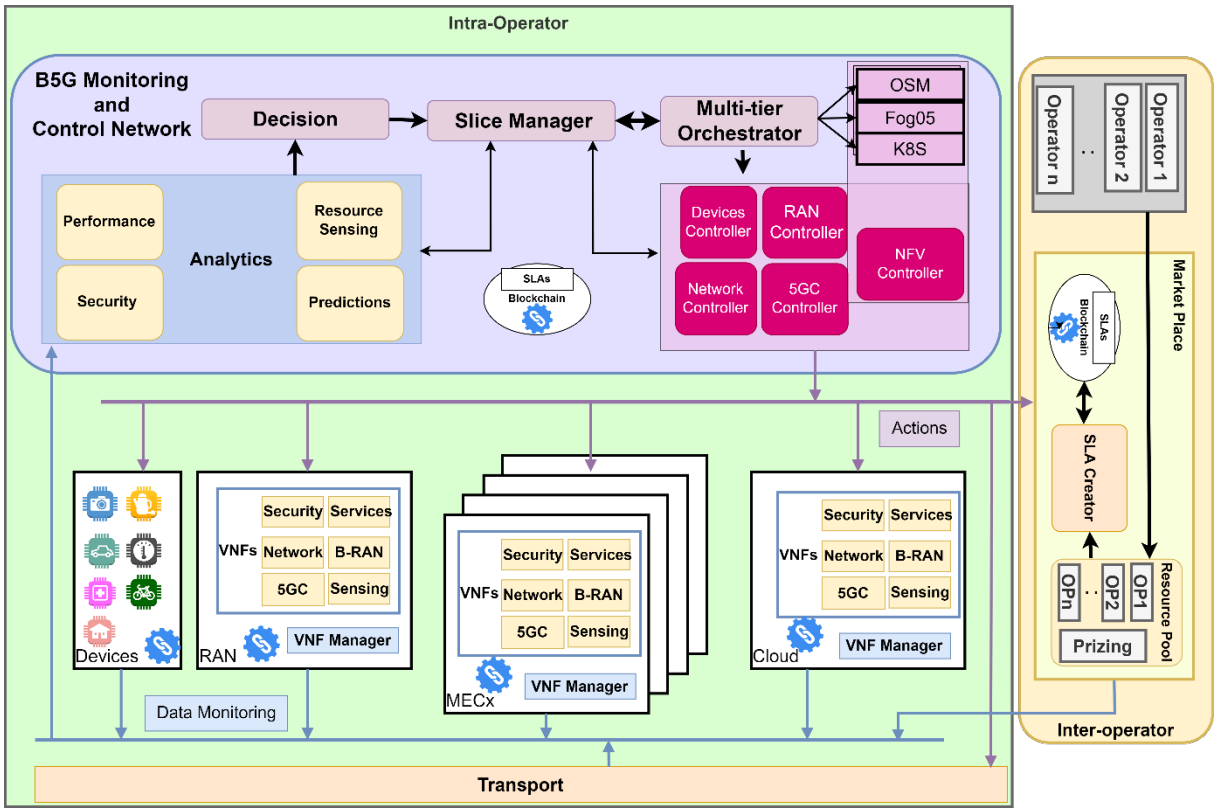

Figure 19 NANCY orchestration architecture

B5G architectures are composed of several domains, RAN, Edge, Transport, Core, Cloud... which underlying technologies and infrastructure are specifically related to the specific domain. Despite the fact that virtualization provides flexibility to the functionality of each of the domains, certain requirements affect directly how the domain must work and cooperate with other domains to maintain the stability of the system and services.

In Figure 19, the NANCY orchestration architecture is proposed, the framework is characterized by:

- A single B5G Monitoring and Control Network Domain: It has a global view of the resources and capabilities of the different underlying domains. The B5G Monitoring and Control Network Domain coordinates the procedures involved in the deployment and management of slices, by enforcing the different actions that should take place on each of the domains to provide connectivity, reliability, security, and QoS. The E2E management domain is fed by the data provided by the domains belonging to the operator, allowing the dynamic correlation of heterogeneous structured data to make accurate decisions, based on different ways of analysis.
- On the other hand, each of the domain's control management interacts directly with the infrastructure to deploy and configure specific assets, serving as an entry point for the Orchestration domain to enforce intents and for the devices to request task offloads. The orchestration will be adapted depending on the availability of orchestration services (OSM, K8S, OpenStack...), the local domain controllers (e.g. RAN Controller) and available technologies.
- Finally, in order to enable inter-operator resource sharing, the virtualised resources will be offered and requested by the different participating operators in a centralized Marketplace which has to be continuously within the orchestration loop in order to be updated with the status of the traded resources in real-time. The transactions record will be maintained in a distributed blockchain which will ensure the trustworthiness and good operation of the whole system.

A detailed description of the overall building blocks of the orchestration architecture presented in Figure 19 is provided as follows:

Control Loop

The orchestration is automated via intents, which means that different steps are defined, namely, *Decision*, *Acting*, *Monitoring,* and *Analysing*, Figure 20 represents the interactions of the steps, whose functionality is as follows:

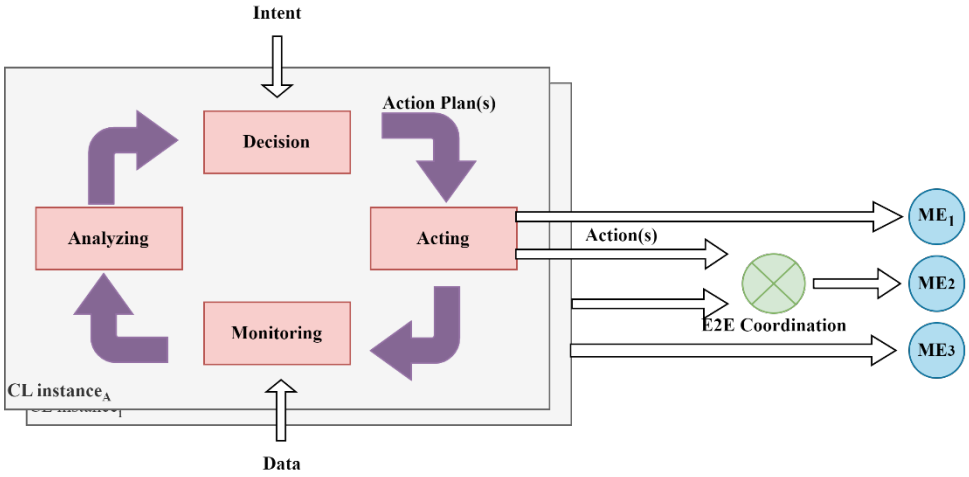

Figure 20 NANCY orchestration control loop

- **Decision**: The Decision step is triggered right after receiving an intent by the SLA derivation procedure from the SLA Manager or by an alert raised due to the analysis performed from the collected data. The AI-based engine will then select the most suitable asset option and location to enforce the received intent.
- **Acting**: Once the intent is created, acting is the process of deploying and configuring, through orchestration, the required services to provide the requirements specified in the SLA.
- **Monitoring**: Every requirement specified in the SLA is monitored to ensure compliance with the SLA, the extracted data from the infrastructure are structured to ease further analysis and correlation.
- **Analyzing**: Monitored data are tracked by the Analysis Engine that yields pattern detection, which feeds the Decision Engine. The pattern detection process involves security anomalies and context prediction (e.g., resource usage or mobility transition).

As a result, all the mentioned Managed Entities (ME) receive precise configurations to adapt to context changes. The functionality that delimits each of the steps could be performed by one or more entities, such as specialized agents that only cover certain technology, or aggregation points that correlate information and can act seamlessly in different domains and technologies.

In this regard, the Task Offloading mechanism will be translated into an intent-based action, where the request will be analysed, in order to determine if the task can be addressed properly (available resources in the involved domains) and the Decision Engine will select the most appropriate solution to the intent (e.g., location and allocated resources), afterwards the Orchestrator will deploy and configure the offloaded task.

Managers and Controllers

These modules support the orchestrator in transferring specific actions to entities and domains which, due to the adoption of a unique type of language and the specificity of the parameters involved, cannot be carried out directly through well-known orchestrators such as OSM} or K8S. The main characteristic of these components is their modularity, allowing the incorporation of new managers and controllers in a seamless and straightforward manner. Thus, enabling the scalability of the system without the need to incorporate changes in the rest of the Control Loop processes.

Decision

The Decision module is based on AI and aims at deducting the best way to resolve requests. These requests can take the form of a SLA or an alert produced by the analytics system. The Decision system will also have extensive knowledge of the existing infrastructure in order to create determined actions to relate requests to specific parts of it. As part of the actions to be executed, the continuous monitoring of the SLA requirements is covered, as well as the corroboration that the recommended action fulfills the objective that motivated it.

Slice Manager

A slice manager is a system that effectively handles shared resources by facilitating the registration, retrieval, and deletion of "partitionable" resources provided by the infrastructure owner. These resources primarily include computing resources, network resources (specifically for the network interfaces of computing resources), and access network resources. The slice manager also plays a crucial role in managing the chunking or partitioning of these resource types. This involves dividing the resources into smaller units, such as compute chunks based on tenants at the Virtualized Infrastructure Manager (VIM) and Network Function Virtualization Orchestrator (NFVO) level, network chunks for isolated networking of tenants, and access chunks based on concepts from the RAN Controller, which represent a subset of wireless interfaces or cells. Additionally, the slice manager is responsible for overseeing the lifecycle of slices, which are collections of the aforementioned resource chunks associated with a specific slice user or vertical. The manager handles various aspects of the slice lifecycle, including slice creation (grouping resource chunks), slice activation (requiring service deployment for slice establishment), and triggering the instantiation of vertical services on slices, along with supporting actions such as DNS provisioning. Moreover, the slice manager utilizes vertical service descriptors from the NFVO catalogue to facilitate the provisioning and management of these services within the slices.

Multi-tier Orchestrator

The multi-level orchestrator is an abstraction of the orchestration capabilities concept that offers high flexibility for action deploying across different domains. It conforms the main piece to allow the execution of actions by logically interconnecting the decision module and other architectural components such as OSM or by directly interacting with controllers to allow configuration or instantiation of particular elements associated with specific functions (e.g., RAN controller or SDN switches). Therefore, this module is in charge of deciding where and how to deal with potential actions, relying on the support provided by the controllers and managers, as well as, coordinating clusters of K8S and OSM placed at different locations.

Analytics

The analytics module receives the monitored data from the infrastructure components and performs a state of the system analysis which is essential to check if the established SLA are being met. Additionally, this module performs a continuous learning of the system, allowing its evolution through ML techniques towards optimized continuous KPIs assurance. The analytics module has four main functionalities, namely, (i) measuring the performance of the system, including the analysis of the requirements established in the SLA, (ii) measuring the use of resources, aiming to maintain an optimal state of the resources needed to guarantee the correct functioning of the system, (iii) detecting anomalies in the behaviour of the users or the system itself, focused on guaranteeing permanent security at different levels of the system, and (iv) establishing predictive models, referring to the previous three points, which allows early action in situations that could put the quality and continuity of the services at risk.

Blockchains

Blockchain agents are responsible for keeping a record of the resource-sharing agreements made between different operators in the SLA format. In addition, for each of the leased resources, the blockchain will also irrefutably certify the resources in each of the domains necessary to meet the SLA; for each of these resources, the price agreed between the operators will be added as well. In essence, the blockchain agents are located in all those places of the infrastructure that can be shared directly or indirectly with other operators, in such a way that they serve to increase the distribution of the blockchain nodes and thus, enhance the reliability of the system.

MarketPlace

This module comprises a centralized location connected to all operators participating in the blockchain system, and therefore able to share and request resources. The marketplace is responsible for registering all requests from the different operators, both to publish resources and to request them. In addition, when resources are requested, the marketplace is responsible for generating the SLA that is forwarded to the operator's orchestrator. Prior to the deployment, the SLA must be analyzed by the operators to check the feasibility before being added to the blockchain. Once the operators have accepted the SLA the deployment procedure starts via orchestration. The marketplace has the following components:

- Resource Pool: It is the place where operators publish and request resources available. The resources are related with a prizing and grouped by operators or type of resource.
- SLA creator: module that receives the request and creates and SLA as a contract between customer operator and provider operators.
- Blockchain: one node of the blockchain in which accepted SLA between operators will be added as a block.

4.4.1. VNFs Orchestration

In NANCY, the orchestration and deployment of VNFs are accomplished with different technologies according to the use case and the domain where these functions must be deployed. According to the ETSI MANO the Slice Manager relies on the NFVO to address all the management requests that require the deployment or reconfiguration of a function or service. The NFVO, in turn, translates this request into a specific action that involves creating, deleting, or moving a VM or container. Specifically, this action is delegated to the VIM (Openstack or Kubernetes). The interconnection between Slice

Manager, NFVO, and orchestration solutions is better depicted in Figure 21 (where the NFVO instance in the Openstack case is realised by the OSM block).

Depending on the domain where the virtual functions are deployed, alternative virtualization solutions might be involved. Specifically in the edge domain, a virtualization technology based on VOSySmonitor can be considered for the deployment of VMs, labeled "VOS VM" in the picture. These VMs are specifically designed for the edge where ARM-based embedded systems can be utilized as energy-efficient servers. VOS VMs are not technically VMs, but rather partitions of the hardware where an OS runs bare metal. This solution yields better efficiency compared to VMs as there's no involvement of a hypervisor layer. From the management perspective, these partitions can be viewed as VMs, which can be deployed and destroyed as needed by the VIM.

On more general Linux-based machines, the performance of containers and virtual machines can be managed by leveraging the SCHED_DEADLINE scheduling class of Linux [200], which allows reserving a portion of the available processing bandwidth to a specific virtual CPU, also guaranteeing a maximum service delay. This also allows for avoiding underutilization of the computing platforms that occurs when using a coarse-grained one-to-one mapping between physical and virtual CPUs, which results in inefficiency in the presence of lightweight network functions. SCHED_DEADLINE was originally implemented to assign only one thread to each reservation; nevertheless, it has been more recently extended to support groups of threads within the same reservation [181], thus becoming an attractive option for Linux-based containers and VMs.

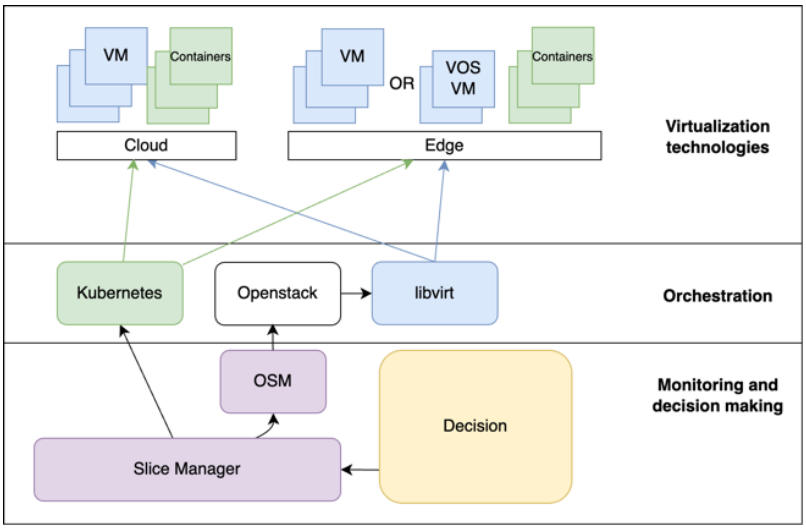

Figure 21 Virtualization technologies overview

4.5. NANCY architectural extensibility

4.5.1. Reconfiguration

The proliferation of mobile Internet and advancements in 5G and beyond technology have resulted in the appearance of new services and an increase in the already substantial volume of data traffic. This places a significant strain on resource management. An efficient and adaptable resource management system is crucial in B-RAN to handle multidimensional resources such as radio, computer, and storage resources. The use of these resources is impacting the efficiency of the network, and the coordination of their administration presents difficulties.

Spectrum resources are the fundamental radio resources in wireless communications, and their availability is becoming more limited due to the growing demands of various applications for low latency and high throughput. B-RAN provides a spectrum-sharing platform that is transparent to all players, including both the spectrum holders and requesters. More precisely, the spectrum-sharing platform may be constructed using nodes that possess ample storage capabilities, such as APs, edge nodes, and UEs. Subsequently, the data on spectrum access and usage may be securely stored in the blockchain, ensuring data immutability, and facilitating the decentralized administration of the spectrum. On this platform, the users submit smart contracts that include information about their spectrum requests, such as bandwidth requirements and power use cycle, to be published on the public spectrum. The access service providers have the ability to verify their local copies of the distributed ledger, ascertain the spectrum allocation plan, and authenticate the smart contract. The exchange of spectrum between the individuals making requests and those responding may be efficiently executed via the use of a smart contract. Subsequently, the spectrum resources associated with the smart contract will be automatically granted authorization to the individual making the request. Furthermore, it should be noted that individuals who possess spectrum prioritize their own interests and are often hesitant to share them because of the detrimental effects caused by co-channel interference. In order to tackle this problem, it is possible to create incentive mechanisms based on blockchain technology. These mechanisms would encourage spectrum holders to share their spectrum and ensure that all parties involved in a transaction adhere exactly to the terms of the transaction. This would ultimately enhance the use of wireless networks' spectrum.

B-RAN includes a vast array of edge devices equipped with computational capabilities. Efficient management and reconfiguration of these computational resources may greatly enhance the performance of the B-RAN. This is accomplished by the AI-orchestration engine of NANCY and the computational offloading and caching mechanisms. Specifically, the resource requests are handled via the smart contract, which includes details such as task computing information, necessary computing resources, time limitations, and rewards. The resource providers who own unused computing resources and are interested in renting them may access the smart contract. They can review the tasks and decide whether to accept the request based on their available resources and predicted costs. A smart contract will be executed after both sides have concluded the match. The computational results and transaction details will be uploaded to the blockchain network for authentication. Additionally, a reputation incentive system may be established inside the smart contract, whereby the reputation of a participant is assessed based on the quality of service and their display of honest conduct. Participants with a better reputation will be given more precedence when it comes to posting or receiving requests. Finally, storage resource management is crucial in B-RAN because the edge nodes may be hesitant to contribute their store resources owing to capacity limitations. In order to address this issue, it is necessary to provide cache incentive mechanisms that encourage the edge nodes to share their cache capacity. Blockchain-based smart contracts are used to provide a decentralized agency mechanism, allowing content providers to compete and cooperate in order to achieve optimal content distribution performance.

4.5.2. Flexibility/Elasticity

The proposed B-RAN architecture enables the integration of customizable smart contracts to facilitate the flexible administration of diverse services. It is designed to be compatible with both existing networks and upcoming applications. For instance, RAN sharing encompasses several scenarios, ranging from the sharing of spectrum to the sharing of infrastructure. These scenarios are anticipated to enhance both spectral and energy efficiency, while also reducing operating and capital expenses.

This scenario integrates the adaptability of O-RAN software and virtualization with the responsiveness and adaptability of closed-loop control. In this respect, the field of network management and orchestration is progressively shifting towards higher degrees of automation and complete closed-loop control. This is facilitated by the simultaneous implementation of developments in AI/ML technology. The purpose of this shift is to establish a structure that effectively promotes dependability, adaptability, durability, and accessibility by using the notion of "continuum orchestration" - which refers to the seamless coordination among devices, edge computing, and cloud computing to manage changes in infrastructure, needs, and failures.

The existing blockchain solutions are plagued by significant processing and packet overhead, as well as restricted scalability. The scalability of the blockchain might act as a bottleneck, restricting the efficiency of decentralized networks based on blockchain technology. However, this pace is considered too sluggish for maintaining highly dynamic wireless networks. Advanced blockchain technologies, like Lightning and Raiden, are anticipated to assist in resolving dynamic networking issues.

In order to effectively use B-RAN on a broad scale in the future, it is crucial to create a blockchain architecture that can be easily expanded and supports the collaboration of several chains. Nevertheless, the interchange of information across various chains is challenging owing to the obstacles that exist among them. This poses a significant difficulty in terms of scalability when implementing the B-RANs. The obstacles that restrict the sharing of information across different chains may be categorized into two main types: On the one hand, it is necessary to validate the transaction status from the previous chain in the current chain using a distributed approach. On the other hand, it is crucial that the total quantity of the token from the preceding chain remains unchanged throughout the cross-chain transaction procedure. Network sharding is a crucial technique for enhancing the scalability of BF-RAN. It serves as an on-chain growth tool. Blockchain sharing involves the partitioning of nodes in a blockchain network into many autonomous shardings. Each sharding operation handles and saves a tiny segment of transactions and network states, respectively. Consequently, the use of numerous shardings enables the simultaneous processing of transactions, thus enhancing the overall throughput of the networks. Nevertheless, when the quantity of shardings escalates, the level of security in the blockchain network diminishes in a linear fashion. Furthermore, blockchain is fundamentally a decentralized method for managing ledgers, which involves the recording and verification of transactions. Within the B-RANs, every node in the network is required to maintain a portion or the whole of this ledger. As the number of transactions grows, this may result in a significant storage strain on the edge nodes, particularly those with low storage capacity.

5. Requirements of the In-lab Testbeds

Italian in-lab testbed

The Italian in-lab testbed hosts the hardware and software components that allow reproducing, in a limited indoor isolated environment, the Usage scenario "Fronthaul network of fixed topology" in two different situations: "Direct connectivity" and "Coordinated Multi-Point (CoMP) connectivity". These usage scenarios are detailed in D2.1.

Furthermore, the Nancy architectural components that are relevant for the Italian in-lab testbed are those linked to the NANCY's project results listed below:

- [R1] B-RAN architecture and attacks modelling
- [R2] Novel trustworthy grant/cell-free cooperative access mechanisms
- [R3] A novel security and privacy toolbox that contains lightweight consensus mechanisms, and decentralized blockchain components
- [R4] Realistic blockchain and attacks models and an experimental validated B-RAN theoretical framework
- [R5.3] Quantum Safety Mechanisms - PQC signature blockchain
- [R6] Smart pricing policies
- [R7] AI-based B-RAN orchestration with slicer instantiator
- [R8] A novel AI virtualiser for underutilized computational and communication resource exploitation
- [R9] Novel self-evolving AI model repository
- [R10] Experimentally driven reinforcement learning optimization of B-RAN
- [R11] Semantic & goal-oriented communication schemes for beyond Shannon excellence
- [R12] An explainable AI framework
- [R14] A computational offloading mechanism with novel resource-aware/ provision scaling mechanisms and novel battery as well as computational capabilities aware offloading policies
- [R15] User-centric caching mechanisms

Based on the above and on the NANCY architecture, the initial architectural components relevant to Italian in-lab testbed that have been identified are:

- Blockchain
- Theoretical modelling verification
- Anomaly detection
- AI-based orchestration
- Task offloading
- Caching mechanisms

The following Figure 23 shows the initial environment, alongside the schematics of the main functionalities, made available in the ITL Italian indoor lab for NANCY's testing activities relevant for Italian in-lab testbed.

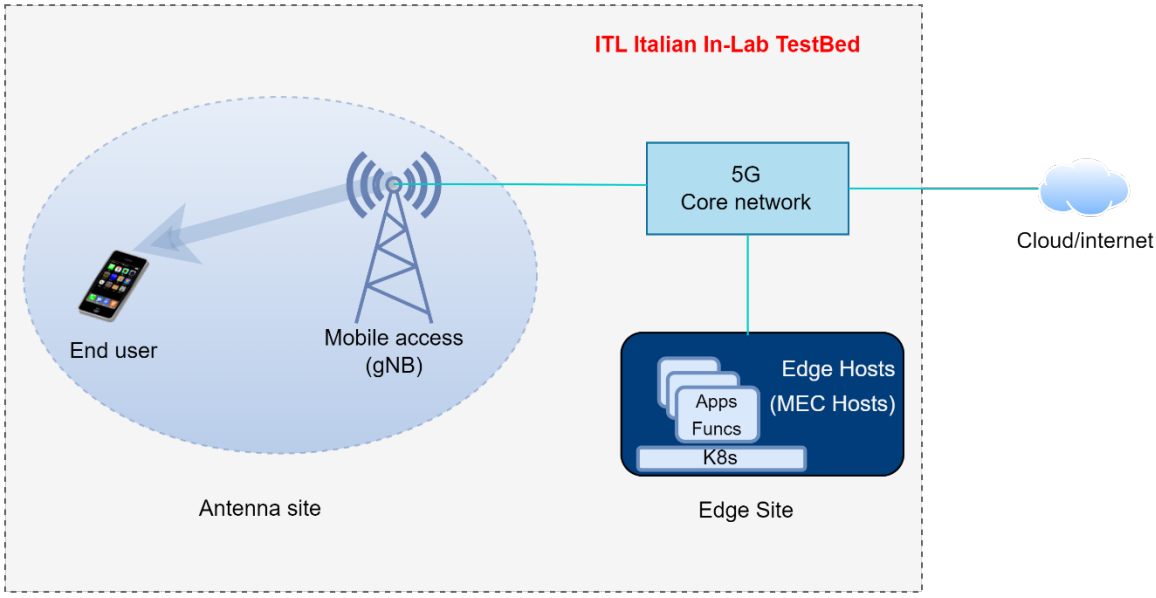

Figure 22 Schematics of the main functionalities of the ITL Italian indoor lab

Taking into account the above, the Antenna site and the Edge site, provided by ITL lab, will host the O-RAN functions together with the required extensions and the architectural components specified and developed by NANCY that are relevant for the Italian in-lab testbed.

Based on these initial requirements, coming from the NANCY architecture definition, D6.4 will specify in detail the architectural design of the testbed, together with the characteristics and the detailed specification of the hardware and software components, and the activities plan for the Italian in-lab testbed.

Greek in-lab testbed

The Greek in-lab testbed will realize the coverage expansion use case. To this end, the following hardware equipment is used for deploying two 5G BSs (i.e., two gNodeBs):

1. Two Ettus Research USRP B210, one acting as the main BS and one as the intermediate BS.
2. Two high-performance laptops for managing the USRPs using the USRP Hardware Driver
    - Intel i7 20-thread & Intel i7 12-thread CPUs
    - 32 & 16 GB of RAM
3. A Quectel RM520N-GL 5G Module is used to connect the intermediate node to the main BS.

Furthermore, a Waveshare 5G Hat, based on the SIM8200A-M2 5G module, is used as user equipment (UE) that connects to the main and intermediate BS. Finally, two programmable sysmocom sysmoISIM-SJA2 subscriber identity modules (SIMs) are programmed to connect and authenticate with the respective BSs.

The hardware components of the testbed are illustrated in Figure 24:

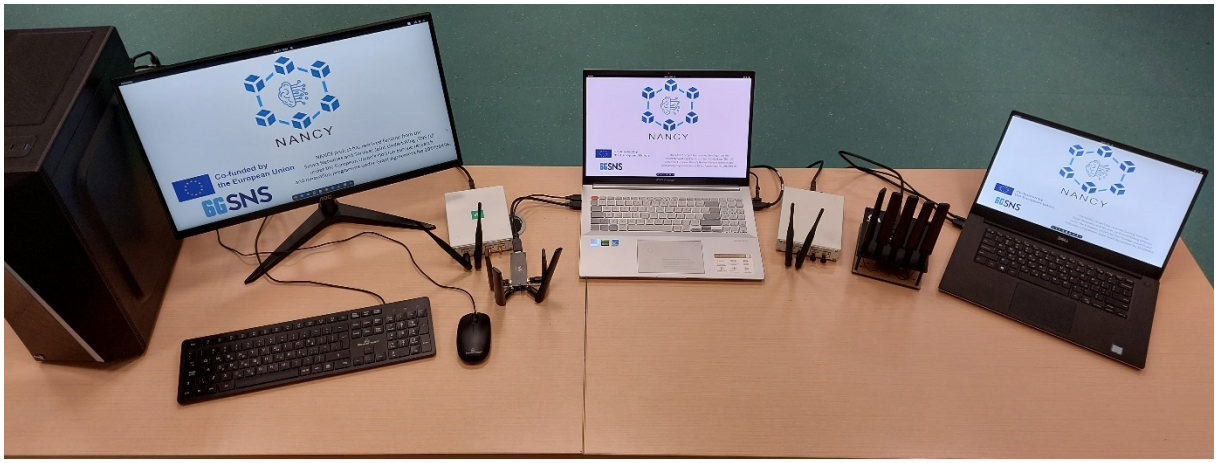

Figure 23 Picture of the testbed equipment

Concerning the software components, Open5GS is used for providing core network functionality, while srsRAN is used for deploying a USRP-based 5G BS. Moreover, FlexRIC is used as a near-real time RAN intelligent controller.

The experiments that will take place in the Greek in-lab testbed involve two scenarios, namely A) a scenario in which the UE is connected directly to the BS, and B) a scenario in which an intermediate node is used to provide coverage to the UE. The two scenarios are depicted in the following Figure 25:

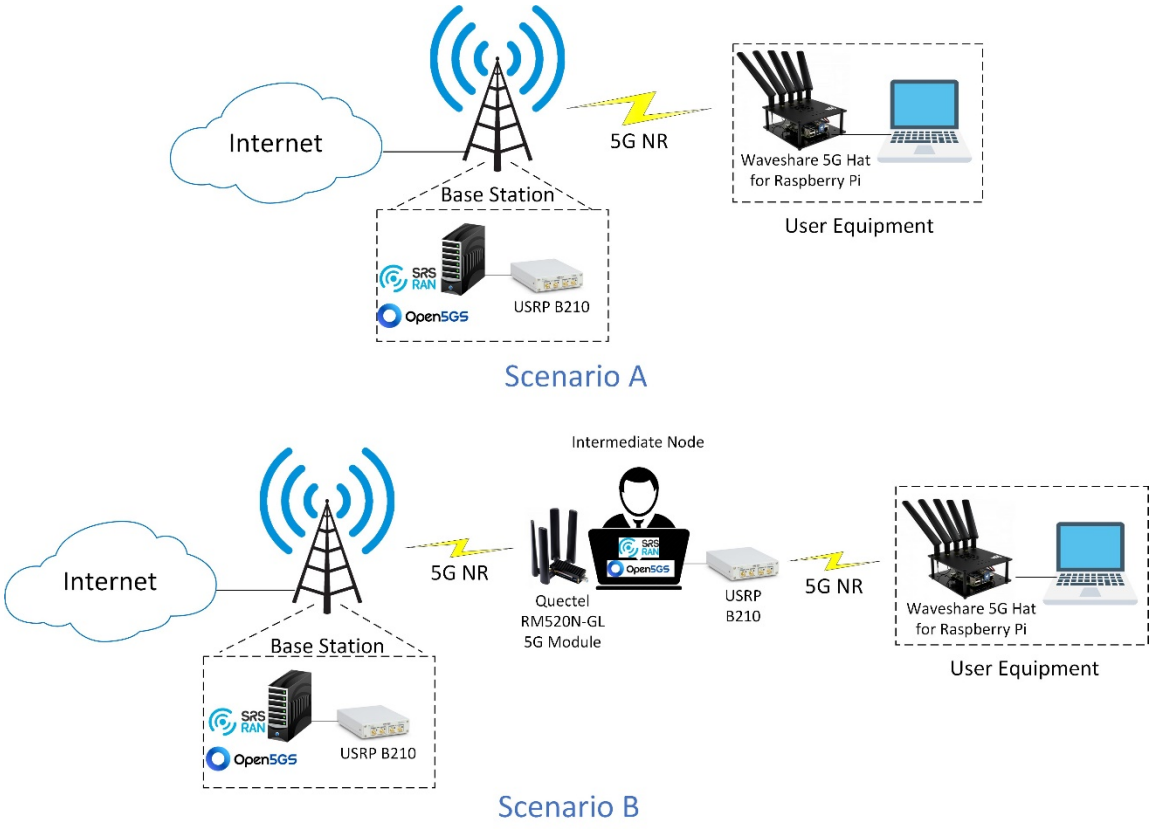

Figure 24 Scenarios A-direct connection to the BS, and B-connection to the BS trough intermediate node

The Greek in-lab testbed is focused on evaluating and validating the NANCY outcomes in coverage expansion scenarios. To this end, a video streaming application will be employed. In general, streaming applications require the transmission of large data volumes with low latency. Therefore, they are

useful assets for evaluating and assessing network performance in real-time. Furthermore, other traffic generation tools, such as the iPerf3 software, will be employed to create additional network traffic.

In more detail, the video streaming application and the traffic generation tools will run in both scenarios and the respective network performance metrics will be assessed. The self-recovery/healing and caching are two aspects that will be also investigated in the Greek in-lab testbed. Specifically, the intermediate node may become unavailable (for instance, when the node is turned off or is compromised due to a cyberattack). In addition, a moving device may need to disconnect from the BS and connect to the intermediate node for energy-efficiency reasons. To ensure a constant stream of video data, the handover between the BS and the intermediate node will have to take place with minimum delay.

Italian Massive IoT Testbed

The Italian Massive IoT testbed aims to demonstrate a fixed topology fronthaul network with direct connectivity. This setup could showcase specific aspects and applications related to IoT Technology.

Fronthaul networks can indeed be deployed in both point-to-point and point-to-multipoint topologies, offering different advantages based on the specific needs of the network and the applications it supports.

The testbed highlights three major areas where NANCY usage scenarios can be demonstrated:

- uRLLC: Ultra Reliable Low Latency Communication
- mMTC: Massive Machine Type Communication (IoT)
- eMBB: Enhanced Mobile Broadband – high speed

mainly validating the functionalities mentioned below.

- Self-healing and self-recovery
- Post Quantum Cryptography (PQC)

The Italian commercial 5G testbed is set to utilize a radio access network with new encryption algorithms for secure traffic delivery, illustrated in the following Figure 26.

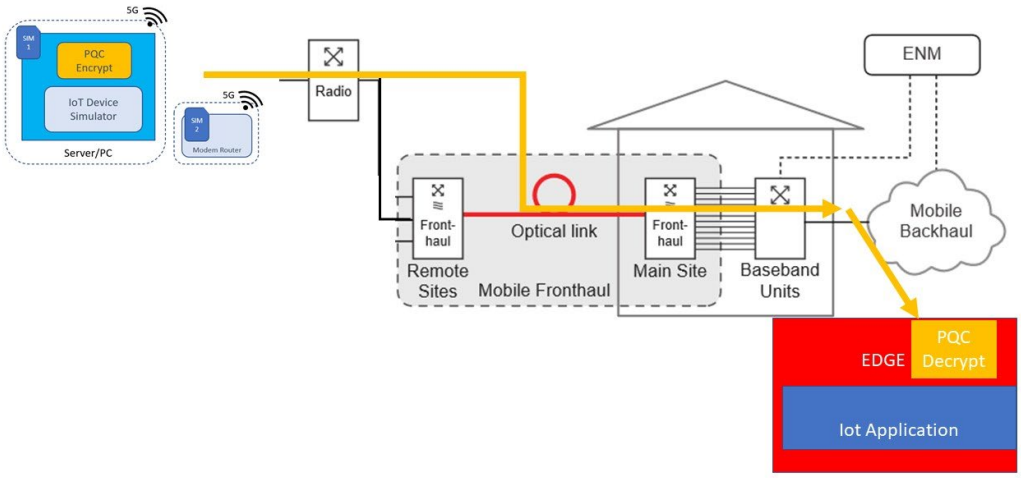

Figure 25 TEI's testbed topology for NANCY

The features mentioned above will be deployed and assessed in the testbed located in Genoa (Italy). This testbed is part of the Italian Ericsson Research and Development organization, and it is a private Network in a controlled environment.

The Ericsson Mobile Network used for the tests will evolve throughout the project: it allows to collect results during the applications integration facilitating the comparison of the performances in the two architectures.

In the first phase, a non-standalone (NSA) architecture is implemented. From the RAN perspective, the Ericsson Radio4422 and the Baseband 6648 will be used. In this scenario, the server for the application will be connected to the EPG and reachable by the client side through a dedicated 5G device. The fronthaul connection is implemented by the optical Ericsson FH6000 family nodes, allowing a fully flexible optical transport solution.

The second phase foresees the usage of a standalone (SA) architecture with the advantages of a full 5G mobile network. From a hardware perspective, the baseband will be the same (BB6648), the fronthaul solution will be based on the same FH6000 family nodes scaling the fronthaul rate at 25Gb (eCPRI), while the radio solution will be evaluated, depending on the phase 1 test results, between a wide range of available nodes. Nowadays AIR6419 and AIR1281 are the two considered solutions: they are already integrated in the laboratory, and they could be the most interesting implementation for the project scope.

6. Requirements of the Outdoor Testbeds

Greek Outdoor Testbed

In the Greek outdoor testbed, the following sub-use cases are expected to be demonstrated.

- Advanced coverage expansion-multi-hop connectivity
- Advanced coverage expansion-Ad-hoc mesh connectivity
- Advanced coverage expansion-Point-to-multipoint connectivity

validating the functionalities mentioned below.

- AI-based B-RAN orchestrator
- Computational offloading and social-aware caching, and
- Smart pricing

The testbed that will be provided by OTE for the implementation of the NANCY Greek outdoor demonstrator is depicted in the following Figure 27:

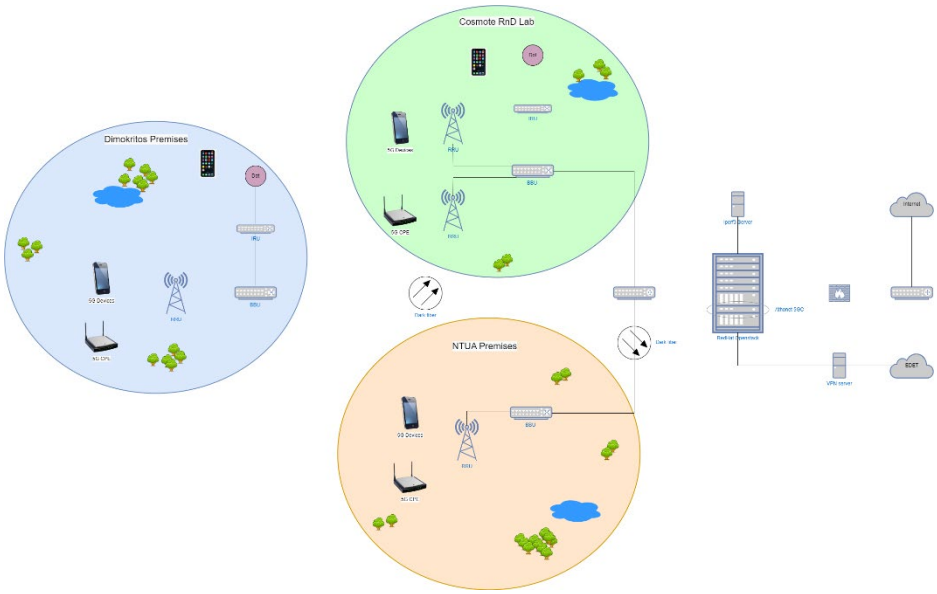

Figure 26 Topology of the OTE's outdoor testbed

To deploy and assess the aforementioned features of the NANCY architecture, this testbed will be located in Athens, under the leadership of OTE and it will use the following equipment:

Mobile network hardware: OTE will provide an ATHONET 5G SA core network, including:

- Two UPFs to emulate edge and core 5G network Data planes
- The following 3GPP Control Plane Network Functions: Access and Mobility Management Function (AMF), Session Management Function (SMF), Authentication Server Function (AUSF), User Data Management (UDM) Function
- Supporting 3GPP interfaces: N1, N2, N3, N4, N6 to be hosted at OTE Cloud facilities
- Network Slicing

The RAN will consist of:

- Ericsson BBU 6630
- Radio Unit 4408
- IRU 8848 + Dot 4479 B78L

The partners-owners that participate in this demonstration will integrate their components in the above-mentioned testbed for the successful implementation of the respective functionalities.

Spanish outdoor testbed

The Spanish outdoor testbed hosts the hardware equipment and the software components that allow reproducing the usage scenario "Advanced connectivity of mobile nodes" (detailed in D2.1). This scenario supports vehicle-to-AP and vehicle-to-vehicle communications.

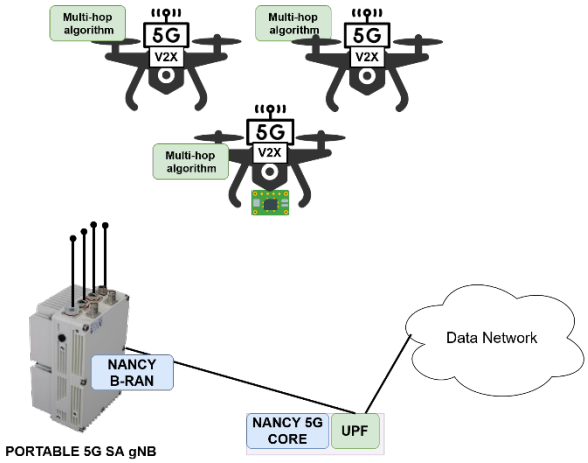

Figure 27 Usage scenario for Advanced connectivity of mobile nodes

For this, we estimate that the following physical communication and computation equipment is necessary:

- 1x Portable 5G gNB
- 3x UAV
- 3x V2X portable modem
- 3x 5G portable modem
- 1x Laboratory 5G gNB+core (optional)
- 3x Raspberry Pi (or similar)

For controlling the networking and computation elements, we estimate that the following software artifacts are needed:

- Development 5G core (free5Gcore v3.2.1)
- Cloud orchestrator (Kubernetes v1.28)
- Virtual infrastructure manager (OpenStack version wallaby)
- SDN controller (ONOS v2.6)
- UAV controller (PIXHAWK)
- Software for multi-hop communications (probably self-developed)

The NANCY architectural components that are relevant for this testbed are those linked to the following project results:

- [R1] B-RAN architecture and attacks modelling
- [R2] Novel trustworthy grant/cell-free cooperative access mechanisms
- [R3] A novel security and privacy toolbox that contains lightweight consensus mechanisms, and decentralized blockchain components
- [R7] AI-based B-RAN orchestration with slicer instantiator
- [R8] A novel AI virtualiser for underutilized computational and communication resource exploitation
- [R13] Next-generation SDN-enabled MEC for autonomous anomaly detection, self-healing and self-recovery
- [R14] A computational offloading mechanism with novel resource-aware provision scaling mechanisms and novel battery as well as computational capabilities aware offloading policies
- [R15] User-centric caching mechanisms

According to the testbed requirements, the initial architectural components that have been identified are:

- Blockchain
- Grant/cell-free cooperative access
- AI-based orchestration
- AI virtualiser
- Task offloading
- Social-aware caching
- Anomaly detection, self-healing and self-recovery
- VNFs orchestration

7. Conclusion and Outlook

This deliverable outlined the novel NANCY architecture, aimed at advancing B5G wireless networks by establishing a ground-breaking framework that integrates blockchain technologies and AI to enhance network security, efficiency, and intelligence. The architecture addresses key gaps in current B5G/6G architectures, particularly in the realms of energy efficiency, security, and intelligent resource management. The analysis of existing 5G/6G state-of-the-art, alongside the integration of insights from both 5GPPP and non-5GPPP projects, will ensure that NANCY remains at the forefront of technological innovation.

Additionally, this deliverable outlines NANCY's ambition to redefine and enhance traditional network architectures, proposing novel solutions like P2P connectivity, mesh networking, and relay-based communications, which are vital for optimizing the network. These optimizations are expected to transform the network into a low-power, distributed, and intelligent framework for wireless networking technology. Furthermore, the NANCY project's holistic approach, combining advancements in blockchain, multi-access edge computing, and AI, paves the way for personalized, multi-tenant, and perpetually protected wireless networks. The project's contributions also extend to establishing a new, experimentally verified network information-theoretic framework, accommodating the unique requirements introduced by these novel technological building blocks.

Bibliography

[1] F. Tschorsch and B. Scheuermann, "Bitcoin and beyond: A technical survey on decentralized digital currencies," *IEEE Commun. Surv. Tutorials,* vol. 18, no. 3, p. 2084–2123, Thirdquarter 2016.

[2] Y. Dai, D. Xu, S. Maharjan, Z. Chen, Q. He and Y. Zhang, "Blockchain and deep reinforcement learning empowered intelligent 5G beyond," *IEEE Network,* vol. 33, no. 3, p. 10–17, May 2019.

[3] Cisco Visual Networking Index, *Cisco visual networking index: Global mobile data traffic forecast update, 2016-2021,* Cisco, White Paper, Feb. 2017.

[4] X. Ling, J. Wang, T. Bouchoucha, B. C. Levy and Z. Ding, "Blockchain radio access network (B-RAN): Towards decentralized secure radio access paradigm," *IEEE Access,* vol. 7, p. 9714–9723, Jan. 2019.

[5] K. Christidis and M. Devetsikiotis, "Blockchains and smart contracts for the Internet of Things," *IEEE Access,* vol. 4, p. 2292–2303, Jan. 2016.

[6] J. Backman, S. Yrjölä, K. Valtanen and O. Mämmelä, "Blockchain network slice broker in 5G: Slice leasing in factory of the future use case," in *Internet Things Bus. Models, Users, Netw.*, Copenhagen, Danmark, Nov. 2017.

[7] "Sunfish Project," [Online]. Available: http://www.sunfishproject.eu/

[8] "SERIOT Prokeject," [Online]. Available: https://seriot-project.eu/

[9] W. Wang, D. T. Hoang, P. Hu, Z. Xiong, D. Niyato, P. Wang, Y. Wen and D. I. Kim, "A survey on consensus mechanisms and mining strategy management in blockchain networks," *IEEE Access,* vol. 7, p. 22328–22370, Jan. 2019.

[10] Y. Dong, J. Cheng, M. J. Hossain and V. C. Leung, "Secure distributed on-device learning networks with byzantine adversaries.," *IEEE Network ,* 2019.

[11] J. Yang, S. He, Y. Xu, L. Chen and J. Ren, "A trusted routing scheme using blockchain and reinforcement learning for wireless sensor networks," *Sensors,* vol. 19, no. 4, p. 970, 2019.

[12] Z. Liu, N. C. Luong, W. Wang, D. Niyato, P. Wang, Y. Liang and D. I. Kim, "A survey on blockchain: A game theoretical perspective," *IEEE Access,* vol. 7, p. 47615–47643, Apr. 2019.

[13] M. B. H. Weiss, K. Werbach, D. C. Sicker and C. E. C. Bastidas, "On the application of blockchains to spectrum management," *IEEE Trans. Cognit. Commun. Networking,* vol. 5, no. 2, p. 193–205, Jan. 2019.

[14] P. Kuo, A. Mourad and J. Ahn, "Potential applicability of distributed ledger to wireless networking technologies," *IEEE Wireless Commun.,* vol. 25, no. 4, p. 4–6, Aug. 2018.

[15] E. D. Pascale, J. McMenamy, I. Macaluso and L. Doyle, *Smart contract SLAs for dense small-cell-as-a-service,* arXiv preprint arXiv: 1703.04502, Mar. 2017.

[16] K. Kotobi and S. G. Bilen, "Secure blockchains for dynamic spectrum access: A decentralized database in moving cognitive radio networks enhances security and user access," *IEEE Veh. Technol. Mag.,* vol. 13, no. 1, p. 32–39, Mar. 2018.

[17] Y. Le, X. Ling, J. Wang and Z. Ding, "Prototype design and test of blockchain radio access network," in *IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, Shanghai, P.R. China, 2019.

[18] U. Challita et al., "When machine learning meets wireless cellular networks: Deployment, challenges, and applications," *IEEE Commun. Mag.,* vol. 58, no. 6, p. 12–18, 2020.

[19] L. Bonati et al., "Intelligence and learning in O-RAN for data-driven NextG cellular networks," *IEEE Commun. Mag.,* vol. 59, no. 10, pp. 21-27, 2021.

[20] M. Polese et al., "Understanding O-RAN: Architecture, Interfaces, Algorithms, Security, and Research Challenges," *IEEE Communications Surveys & Tutorials,* vol. 25, no. 3, pp. 1376-1411, 2023.

[21] S. Teral, “RIC as the next generation SON for Open RAN and more,” *LightCounting,* 2021.

[22] A. Checko et al., “Cloud RAN for mobile networks - A technology overview,” *IEEE Commun. Surveys Tuts.,* vol. 17, no. 1, p. 405–426, 2015.

[23] D. E. C. o. Excellence, “The open future of radio access networks.,” 2021. [Online]. Available: https://www2.deloitte.com/content/dam/Deloitte/pt/Documents/technology-media-telecommunications/TEE/The-Open-Future-of-Radio-Ac.

[24] 3GPP, “Study on new radio access technology: Radio access architecture and interfaces, version 14.0.0,” 2017. [Online]. Available: http://www.3gpp.org/DynaReport/38801.htm.

[25] “O-RAN architecture description 5.00,” O-RAN Working Group 1, 2021. [Online]. Available: https://orandownloadsweb.azurewebsites.net/specifications

[26] 3GPP, “NG-RAN Architecture Description, Version 17.0.0, 3GPP Standard (TS) 38.401,” 2022. [Online]. Available: http://www.3gpp.org/DynaReport/38401.htm.

[27] F. W. Murti et al., “An optimal deployment framework for multi-cloud virtualized radio access networks,” *IEEE Trans. Wireless Commun.,* vol. 20, no. 4, p. 2251–2265, 2021.

[28] 3GPP, “NR Radio Link Control (RLC) Protocol Specification, Version 15.0.0, 3GPP Standard (TS) 38.322,” 2018. [Online]. Available: http://www.3gpp.org/DynaReport/38322.htm.

[29] 3GPP, “NR Medium Access Control (MAC) Protocol Specification, Version 15.0.0 3GPP Standard (TS) 38.321,” 2018. [Online]. Available: http://www.3gpp.org/DynaReport/38321.htm .

[30] 3GPP, “NR Physical Layer; General Description, Version 15.0.0, 3GPP Standard (TS) 38.201,” 2018. [Online]. Available: http://www.3gpp. org/DynaReport/38201.htm.

[31] 3GPP, “NR Radio Resource Control (RRC); Protocol Specification, Version 15.0.0, 3GPP Standard (TS) 38.331,” 2018. [Online]. Available: http://www.3gpp.org/DynaReport/38331.htm.

[32] 3GPP, “Evolved Universal Terrestrial Radio Access (E-UTRA) and NR; Service Data Adaptation Protocol (SDAP) Specification, Version 17.0.0, 3GPP Standard (TS) 37.324,” 2022. [Online]. Available: http://www.3gpp.org/DynaReport/37324.htm.

[33] 3GPP, “NR Packet Data Convergence Protocol (PDCP) Specification, Version 15.0.0, 3GPP Standard (TS) 38.323,” 2018. [Online]. Available: http://www.3gpp.org/DynaReport/38323.htm.

[34] “O-RAN AI/ML workflow description and requirements 1.03,” O-RAN Working Group 2, 2021. [Online]. Available: https://orandownloadsweb.azurewebsites.net/specifications

[35] 3GPP, “Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall Description; Stage 2, Version 14.2.0, 3GPP Standard (TS) 36.300,” 2017. [Online]. Available: http://www.3gpp.org/DynaReport/36300.htm.

[36] T. O’Shea et al., “An introduction to deep learning for the physical layer,” *IEEE Trans. Cogn. Commun. Netw.,* vol. 3, no. 4, p. 563–575, 2017.

[37] “O-RAN cloud architecture and deployment scenarios for O-RAN virtualized RAN 2.02,” O-RAN Working Group 6, 2021. [Online]. Available: https://orandownloadsweb.azurewebsites.net/specifications

[38] “O-RAN acceleration abstraction layer FEC profiles 1.0,” O-RAN Working Group 6, 2021. [Online]. Available: https://orandownloadsweb.azurewebsites.net/specifications

[39] “O-RAN acceleration abstraction layer general aspects an principles 1.01,” O-RAN Working Group 6, 2021. [Online]. Available: https://orandownloadsweb.azurewebsites.net/specifications

[40] A. Nasrallah et al., “Ultra-low latency (ULL) networks: The IEEE TSN and IETF DetNet standards and related 5G ULL research,” *IEEE Commun. Surveys Tuts.,* vol. 21, no. 1, p. 88–145, 2019.

[41] A. Kelkar et al., “NVIDIA aerial GPU hosted AI-on-5G,” in *4th 5G World Forum (5GWF)*, 2021.

[42] G. Garcia-Aviles at al., “Nuberu: Reliable RAN virtualization in shared platforms,” in *27th Annu. Int. Conf. Mobile Comput. Netw.*, 2021.

[43] J. S. Panchal et al., “Enabling and scaling of URLLC verticals on 5G vRAN running on COTS hardware,” *IEEE Commun. Mag.,* vol. 59, no. 9, pp. 105-111, 2021.

[44] E. A. Papatheofanous et al., “LDPC hardware acceleration in 5G open radio access network platforms,” *IEEE Access,* vol. 9, p. 152960–152971, 2021.

[45] D. S. Sabella et al., “Energy efficiency benefits of RAN-as-a-service concept for a cloud-based 5G mobile network infrastructure,” *IEEE Access,* vol. 2, p. 1586–1597, 2014.

[46] D. López-Pérez et al., “A survey on 5G radio access network energy efficiency: Massive MIMO, lean carrier design, sleep modes, and machine learning,” *IEEE Commun. Surveys Tuts.,* vol. 24, no. 1, p. 653–697, 2022.

[47] L. Bonati et al., “Open, programmable, and virtualized 5G networks: State-of-the-art and the road ahead,”,” *Comput. Netw.,* vol. 182, pp. 1-28, 2020.

[48] S. D’Oro et al., “dApps: Distributed applications for real-time inference and control in O-RAN,” *IEEE Commun. Mag.,* vol. 60, no. 11, pp. 52-58, 2022.

[49] X. Foukas, B. Radunovic, M. Balkwill and Z. Lai, “Taking 5G RAN Analytics and Control to a New Level,” *Proceedings of the 29th Annual International Conference on Mobile Computing and Networking,* pp. 1-16, 2023.

[50] E. Björnson et al., “Scalable cell-free massive MIMO systems,” *IEEE Trans. Commun.,* vol. 68, no. 7, p. 4247–4261, 2020.

[51] H. Q. Ngo, A. Ashikhmin, H. Yang, E. . G. Larsson and T. L. Marzetta, “Cell-Free Massive MIMO Versus Small Cells,” *IEEE Transactions on Wireless Communications,* vol. 16, pp. 1834-1850, 2017.

[52] T. Pamuklu, S. Mollahasani and M. Erol-Kantarci, “Energy-Efficient and Delay-Guaranteed Joint Resource Allocation and DU Selection in O-RAN,” *2021 IEEE 4th 5G World Forum (5GWF),* pp. 99-104, 2021.

[53] L. Giupponi and W. Francesc, “Blockchain-enabled network sharing for O-RAN in 5G and beyond,” *IEEE Network,* vol. 36, no. 4, pp. 218-225, 2022.

[54] H. Xu, L. Zhang and Y. Sun, “BE-RAN: Blockchain-enabled open RAN with decentralized identity management and privacy-preserving communication,” *arXiv preprint arXiv:2101.10856,* 2021.

[55] S. Velliangiri, R. Manoharan, S. Ramachandran and V. Rajasekar, “Blockchain based privacy preserving framework for emerging 6G wireless communications,” *IEEE Transactions on Industrial Informatics,* vol. 18, no. 7, pp. 4868--4874, 2021.

[56] M. Enayati, S. S. Lekshmi, T. Toby, M. Prabhu, K. Rahul, S. Parvathy and S. Ponnekanti, “Blockchain-based location sharing in 5G Open RAN infrastructure for sustainable communities,” in *Intelligent Sustainable Systems: Selected Papers of WorldS4 2021, Volume 1*, Springer, 2022, pp. 571-585.

[57] D. Tapscott and A. Tapscott, Blockchain revolution: how the technology behind bitcoin is changing money, business, and the world, Penguin, 2016.

[58] X. Ling, J. Wang, T. Bouchoucha, B. C. Levy and Z. Ding, “Blockchain Radio Access Network (B-RAN): Towards Decentralized Secure Radio Access Paradigm,” *IEEE access,* vol. 7, pp. 9714-9723, 2019.

[59] B. Dai and W. Yu, “Sparse beamforming and user-centric clustering for downlink cloud radio access network,” *IEEE Access,* vol. 2, pp. 1326-1339, 2014.

[60] F. Wilhelmi and L. Giupponi, “On the performance of blockchain-enabled RAN-as-a-service in beyond 5G networks,” in *2021 IEEE Global Communications Conference (GLOBECOM)*, 2021.

[61] J. Wang, X. Ling, Y. Le, Y. Huang and X. You, “Blockchain-enabled wireless communications: a new paradigm towards 6G,” *National science review,* vol. 8, no. 9, 2021.

[62] N. Herbaut and N. Negru, “A model for collaborative blockchain-based video delivery relying on advanced network services chains,” *IEEE Communications Magazine,* vol. 55, no. 9, pp. 70-76, 2017.

[63] J. Kang, R. Yu, X. Huang, S. Maharjan, Y. Zhang and E. Hossain, "Enabling localized peer-to-peer electricity trading among plug-in hybrid electric vehicles using consortium blockchains," *IEEE transactions on industrial informatics,* vol. 13, no. 6, pp. 3154-3164, 2017.

[64] R. Pass and E. Shi, "Fruitchains: A fair blockchain," in *Proceedings of the ACM symposium on principles of distributed computing*, 2017.

[65] A. Dogra, R. K. Jha and S. Jain, "A survey on beyond 5G network with the advent of 6G: Architecture and emerging technologies," *IEEE Access,* vol. 9, pp. 67512-67547, 2020.

[66] C. S. You, J. S. Yeom and B. C. Jung, "Performance analysis of cooperative low-power wide-area network for energy-efficient B5G systems," *Electronics,* vol. 9, no. 4, p. 680, 2020.

[67] Z. Yang, Q. Zhang and Z. Niu, "Throughput improvement by joint relay selection and link scheduling in relay-assisted cellular networks," *IEEE Transactions on Vehicular Technology,* vol. 61, no. 6, pp. 2824-2835, 2012.

[68] M. Shalaby, M. Shahein, M. Shokair and A. Benaya, "The Cell-Free Networks Enhancement by Relays Implementation," *International Journal of Telecommunications,* vol. 3, no. 1, pp. 1-11, 2023.

[69] S. Ranjan, P. Jha, P. Chaporkar and A. Karandikar, "A novel architecture for multihop relaying in 3GPP LTE and 5G networks," in *2019 IEEE Conference on Standards for Communications and Networking (CSCN)*, 2019.

[70] T. Kerdoncuff, T. Galezowski and X. Lagrange, "Mobile relay for lte: Proof of concept and performance measurements," in *2018 IEEE 87th vehicular technology conference (VTC Spring)*, 2018.

[71] Q. Li, A. Charaf, N. Gresset and H. Bonneville, "Radio resource management in next-generation railway system with heterogeneous multi-hop relaying deployment," in *Communication Technologies for Vehicles: 16th International Workshop, Nets4Cars/Nets4Trains/Nets4Aircraft 2021, Madrid, Spain, November 16--17, 2021, Revised Selected Papers 16*, Madrid, Springer, 2021, pp. 59-70.

[72] S. Duan and H. Zhang, "Recent progress on BFT in the era of blockchains," *National Science Review,* vol. 9, no. 10, 2022.

[73] N. Amiet, "Blockchain vulnerabilities in practice," *Digital Threats: Research and Practice,* vol. 2, no. 2, pp. 1-7, 2021.

[74] H. Foundation, "Hyperledger Foundation," Hyperledger, [Online]. Available: https://hyperledger-fabric.readthedocs.io/en/release-2.2/channels.html

[75] D. Ravi, S. Ramachandran, R. Vignesh, V. R. Falmari and M. Brindha, "Privacy preserving transparent supply chain management through Hyperledger Fabric," *Blockchain: Research and Applications,* vol. 3, no. 2, p. 100072, 2022.

[76] M. Vukoli'c, "Rethinking permissioned blockchains," in *Proceedings of the ACM workshop on blockchain, cryptocurrencies and contracts*, 2017, pp. 3-7.

[77] P. Shah, D. Forester, M. Berberich, C. Raspe and H. Mueller, "Blockchain technology: Data privacy issues and potential mitigation strategies," *Practical Law,* 2019.

[78] L. Kuhring, Z. Istvan, A. Sorniotti and M. Vukolic, "StreamChain: Building a Low-Latency Permissioned Blockchain For Enterprise Use-Cases," in *2021 IEEE International Conference on Blockchain (Blockchain)*, IEEE, 2021, pp. 130-139.

[79] Q. Nasir, I. A. Qasse, M. Abu Talib, A. B. Nassif and others, "Performance analysis of hyperledger fabric platforms," *Security and Communication Networks,* vol. 2018, 2018.

[80] P. Thakkar, S. Nathan and B. Viswanathan, "Performance benchmarking and optimizing hyperledger fabric blockchain platform," in *2018 IEEE 26th international symposium on modeling, analysis, and simulation of computer and telecommunication systems (MASCOTS)*, IEEE, 2018, pp. 264-276.

[81] S. Rusch, "High-Performance Consensus Mechanisms for Blockchains," in *12th EuroSys Doctoral Workshop (EuroDW'18)*, 2018.

[82] "Energy Efficiency of Blockchain Technologies. The EU Blockchain Observatory and Forum team," [Online]. Available: https://www.eublockchainforum.eu/sites/default/files/reports/Energy%20Efficiency%20of%20Blockchain%20Technologies_1.pdf

[83] H. Foundation, "Hyperledger," Hyperledger Foundation, 2015. [Online]. Available: https://www.hyperledger.org/projects/fabric

[84] X. Ling, Y. Le, J. Wang, Z. Ding and X. Gao, "Practical modeling and analysis of blockchain radio access network," *IEEE Transactions on Communications,* vol. 69, no. 2, pp. 1021-1037, 2020.

[85] S. Fluhrer, “Reassessing Grover's Algorithm,” *Cryptology ePrint Archive,* 2017.

[86] C. Zalka, “Grover’s quantum searching algorithm is optimal,” *Physical Review A,* vol. 60, no. 4, p. 2746, 1999.

[87] V. Gheorghiu and M. Mosca, “Benchmarking the quantum cryptanalysis of symmetric, public-key and hash-based cryptographic schemes,” *arXiv preprint arXiv:1902.02332,* 2019.

[88] “NIST Computer Security Resource Center,” National Institute of Standards and Technology, 3 January 2017. [Online]. Available: https://csrc.nist.gov/Projects/post-quantum-cryptography

[89] Y. Cao, Y. Zhao, Q. Wang, J. Zhang, S. . X. Ng and L. Hanzo, “The Evolution of Quantum Key Distribution Networks: On the Road to the Qinternet,” *IEEE Communications Surveys & Tutorials,* vol. 24, no. 2, pp. 839 - 894, 2022.

[90] S. Pirandola, U. L. Andersen, L. Banchi, M. Berta, M. Bunandar, R. Colbeck, D. Englund, T. Gehring, L. Cosmo and C. Ottaviani, “Advances in quantum cryptography,” *Advances in optics and photonics,* vol. 12, pp. 1012--1236, 2020.

[91] M. Mehic, M. Niemiec, S. Rass, J. Ma, M. Peev, . A. Aguado, V. Martin, S. Schauer, A. Poppe and C. Pacher, “Quantum key distribution: a networking perspective,” *ACM Computing Surveys (CSUR),* vol. 53, no. 5, pp. 1-41, 2020.

[92] F. Cavaliere, E. Prati, L. Poti, I. Muhammad and T. Catuogno, “Secure quantum communication technologies and systems: From labs to markets,” *Quantum Reports,* pp. 80-106, 2020.

[93] A. K. Fedorov, E. O. Kiktenko and A. I. Lvovsky, “Quantum computers put blockchain security at risk,” *Nature,* 2018.

[94] C.-Y. Li, X.-B. Chen, Y.-L. Chen, Y.-Y. Hou and J. Li, “A new lattice-based signature scheme in post-quantum blockchain network,” *IEEE Access,* vol. 7, pp. 2026-2023, 2018.

[95] . T. M. Fernandez-Carames and . P. Fraga-Lamas, “Towards post-quantum blockchain: A review on blockchain cryptography resistant to quantum computing attacks,” *IEEE access,* vol. 8, pp. 21091-21116, 2020.

[96] Y.-L. Gao, X.-B. Chen, Y.-L. Chen, . Y. Sun, . X.-X. Niu and . Y.-X. Yang, “A secure cryptocurrency scheme based on post-quantum blockchain,” *IEEE access,* vol. 6, pp. 27205-27213, 2018.

[97] E. O. Kiktenko, N. O. Pozhar, M. N. Anufriev, A. . S. Trushechkin, R. R. Yunusov, Y. . V. Kurochkin, A. Lvovsky and . A. K. Fedorov, “Quantum-secured blockchain,” *Quantum Science and Technology,* vol. 3, 2018.

[98] X. Sun, . M. Sopek, Q. Wang and . P. Kulicki, “Towards quantum-secured permissioned blockchain: Signature, consensus, and logic,” *Entropy,* vol. 21, p. 887, 2019.

[99] F. Xu, . X. Ma, Q. Zhang, . H.-K. Lo and . J.-W. Pan, “Secure quantum key distribution with realistic devices,” *Reviews of Modern Physics,* vol. 92, 2020.

[100] CORDIS, “Cordis,” Cordis, [Online]. Available: https://cordis.europa.eu/project/id/101021936

[101] E. U. A. f. Cybersecurity, “Enisa,” European Union Agency for Cybersecurity, 2005. [Online]. Available: https://www.enisa.europa.eu/

[102] ANSSI, “ANSSI,” ANSSI, [Online]. Available: https://www.ssi.gouv.fr/en/.

[103] “BSI,” Federal Office for Information Security , [Online]. Available: https://www.bsi.bund.de/EN/Home/home_node.html

[104] S. Hosny, F. Alotaibi, H. El Gamal and A. Eryilmaz, “Towards a mobile content marketplace,” in *2015 IEEE 16th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, IEEE, 2015, pp. 675-679.

[105] Y. Jiao, P. Wang, D. Niyato and Z. Xiong, “Social welfare maximization auction in edge computing resource allocation for mobile blockchain,” in *2018 IEEE international conference on communications (ICC)*, IEEE, 2018, pp. 1-6.

[106] K. Liu, X. Qiu, W. Chen, X. Chen and Z. Zheng, “Optimal pricing mechanism for data market in blockchain-enhanced internet of things,” *IEEE Internet of Things Journal,* vol. 6, no. 6, pp. 9748-9761, 2019.

[107] M. Casazza, M. Bouet and S. Secci, “Availability-driven NFV orchestration,” *Computer Networks,* vol. 155, pp. 47-61, 2019.

[108] T. Cucinotta, L. Abeni, M. Marinoni, R. Mancini, and C. Vitucci, “Strong temporal isolation among containers in openstack for NFV services,” *IEEE Transactions on Cloud Computing,* vol. 11, no. 1, 2023.

[109] E. Casalicchio, “Container orchestration: A survey,” *Systems Modeling: Methodologies and Tools,* pp. 221-235, 2019.

[110] ETSI TS 128 530, "5G; Management and orchestration; Concepts, use cases and requirements," 2021. [Online]. Available: https://www.etsi.org/deliver/etsi_ts/128500_128599/128530/16.04.00_60/ts_128530v160400p.pdf

[111] Q. Liu, N. Choi and T. Han, “Deep Reinforcement Learning for End-to-End Network Slicing: Challenges and Solutions,” *IEEE Network,* vol. 37, no. 2, pp. 222-228, 2023.

[112] N. Salhab, R. Langar and R. Rahim, “5G network slices resource orchestration using Machine Learning techniques,” *Computer Networks,* vol. 188, p. 107829, 2021.

[113] B. Brik, K. Boutiba and A. Ksentini, “Deep learning for B5G open radio access network: Evolution, survey, case studies, and challenges,” *IEEE Open Journal of the Communications Society,* vol. 3, pp. 228-259, 2022.

[114] S. Saxena and K. M. Sivalingam, “DRL-Based Slice Admission Using Overbooking in 5G Networks,” *IEEE Open Journal of the Communications Society,* vol. 4, pp. 29-45, 2022.

[115] N. Hammami and K. K. Nguyen, “On-policy vs. off-policy deep reinforcement learning for resource allocation in open radio access network,” in *2022 IEEE Wireless Communications and Networking Conference (WCNC)*, IEEE, 2022, pp. 1461-1466.

[116] E. T. S. Institute, “ETSI,” European Telecommunications Standards Institute, [Online]. Available: https://www.etsi.org/

[117] European Telecommunications Standards Institute (ETSI), “Zero touch network & Service Management (ZSM),” [Online]. Available: https://www.etsi.org/technologies/zero-touch-network-service-management

[118] European Telecommunications Standards Institute (ETSI), “Experiential Networked Intelligence (ENI),” [Online]. Available: https://www.etsi.org/technologies/experiential-networked-intelligence

[119] J. Diaz-de-Arcaya, A. I. Torre-Bastida, G. Zarate, R. Minon and A. Almeida, “A Joint Study of the Challenges, Opportunities, and Roadmap of MLOps and AIOps: A Systematic Survey,” *ACM Computing Surveys,* 2023.

[120] T. Zhang, M. Hemmatpour, S. Mishra, L. Linguaglossa, D. Zhang, C. S. Chen, M. Mellia and A. Aghasaryan, “Operationalizing AI in Future Networks: A Bird's Eye View from the System Perspective,” *arXiv preprint arXiv:2303.04073,* 2023.

[121] P. Li, J. Thomas, X. Wang, A. Khalil, A. Ahmad, R. Inacio, S. Kapoor, A. Parekh, A. Doufexi, A. Shojaeifard and others, “Rlops: Development life-cycle of reinforcement learning aided open ran,” *IEEE Access,* vol. 10, p. 113808113826, 2022.

[122] G. Samaras, V. Theodorou, D. Laskaratos, N. Psaromanolakis, M. Mertiri and A. Valantasis, “QMP: A Cloud-native MLOps Automation Platform for Zero-Touch Service Assurance in 5G Systems,” in *2022 IEEE International Mediterranean Conference on Communications and Networking (MeditCom)*, IEEE, 2022, pp. 86-89.

[123] PREDICT-6G, “PREDICT-6G,” [Online]. Available: https://predict-6g.eu/

[124] RISE-6G, “RISE-6G,” [Online]. Available: https://rise-6g.eu/

[125] 5G-IANA, “5G-IANA,” [Online]. Available: https://www.5g-iana.eu/

[126] 5G CLARITY, “5G CLARITY,” [Online]. Available: https://5gclarity.com/

[127] ARIADNE, “ARIADNE,” [Online]. Available: https://www.ict-ariadne.eu/

[128] AI@EDGE, “AI@EDGE,” [Online]. Available: https://aiatedge.eu/

[129] C. Fortuna, D. Mušić, G. Cerar, A. Čampa, P. Kapsalis, and M. Mohorčič, “On-Premise Artificial Intelligence as a Service for Small and Medium Size Setups,” in *Advances in Engineering and Information Science Toward Smart City and Beyond*, Springer, 2023, pp. 53-73.

[130] G. Cerar, B. Bertalanič, M. Mohorčič, M. Grobelnik, and C. Fortuna, “Feature Management for Machine Learning Operation Pipelines in AI Native Networks,” in *2023 International Balkan Conference on Communications and Networking (BalkanCom)*, IEEE, 2023, pp. 1-5.

[131] Y. Bengio, A. Courville and P. Vincent, “Representation learning: A review and new perspectives,” *IEEE transactions on pattern analysis and machine intelligence,* vol. 35, no. 8, pp. 1798-1828, 2013.

[132] J. Kaplan, S. McCandlish, T. Henighan, T. B. Brown, B. Chess, R. Child, S. Gray, A. Radford, J. Wu and D. Amodei, "Scaling laws for neural language models," *arXiv preprint arXiv:2001.08361,* 2020.

[133] G. Cerar and J. Hribar, "Machine Learning Operations Model Store: Optimizing Model Selection for AI as a Service," in *2023 International Balkan Conference on Communications and Networking (BalkanCom)*, Istambul, IEEE, 2023, pp. 1-5.

[134] S. Masoudnia and R. Ebrahimpour, "Mixture of experts: a literature survey," *Artificial Intelligence Review,* vol. 42, pp. 275-293, 2014.

[135] 5GZORRO, [Online]. Available: https://www.5gzorro.eu/

[136] European Commission, "Performance optimization and edge computing orchestration for enhanced experience and Quality of Service - Pledger," [Online]. Available: https://cordis.europa.eu/project/id/871536

[137] M. Barahman, L. M. Correia and L. S. Ferreira, "A Real-time QoS-Demand-Aware Computational Resource Sharing Approach in C-RAN," in *European Conference on Networks and Communications (EuCNC)*, Dubrovnik, Croatia, 2020.

[138] M. A. Khan, M. A. Shah, F. Z. Raja and H. A. Khattak, "A Novel Technique of Dynamic Resource Allocation in Software Defined Network," in *15th International Conference on Emerging Technologies (ICET)*, Peshawar, Pakistan, 2019.

[139] Z. Li, Y. Liu, R. Xin, L. Gao, X. Ding and Y. Hu, "A Dynamic Game Model for Resource Allocation in Fog Computing for Ubiquitous Smart Grid," in *28th Wireless and Optical Communications Conference (WOCC)*, Beijing, China, 2019.

[140] S. Nath and J. Wu, " Deep reinforcement learning for dynamic computation offloading and resource allocation in cache-assisted mobile edge computing systems," *Intelligent and Converged Networks,* vol. 1, no. 2, pp. 181-198, 2020.

[141] J. Wang, S. Zhang and G. Qiu, "Dynamical Resource Allocation for Emergency Communication Systems," in *IEEE 13th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON),*, Vancouver, Canada, 2022.

[142] S. Li and Q. Hu, "Dynamic Resource Optimization Allocation for 5G Network Slices Under Multiple Scenarios," in *Chinese Control And Decision Conference (CCDC)* , Hefei, China, 2020.

[143] P. Popovski, O. Simeone, F. Boccardi, D. Gunduz and O. Sahin, "Semantic-effectiveness filtering and control for post-5G wireless connectivity," *Journal of the Indian Institute of Science,* vol. 100, pp. 435-443, 2020.

[144] J. Oueis and E. C. Strinati, "Uplink traffic in future mobile networks: Pulling the alarm," in *Cognitive Radio Oriented Wireless Networks: 11th International Conference, CROWNCOM 2016, Grenoble, France, May 30-June 1, 2016, Proceedings*, Springer, 2016, pp. 583-593.

[145] E. C. Strinati, T. Haustein, M. Maman, W. Keusgen, S. Wittig, M. Schmieder, S. Barbarossa, M. Merluzzi, H. Klessig, F. Giust and others, "Beyond 5G private networks: the 5g CONNI perspective," in *2020 IEEE Globecom Workshops (GC Wkshps)*, IEEE, 2020, pp. 1-6.

[146] R. Li and others, "Towards a new internet for the year 2030 and beyond," in *Proc. 3rd Annu. ITU IMT-2020/5G Workshop Demo Day*, 2018, pp. 1-21.

[147] G. Berardinelli, N. H. Mahmood, I. Rodriguez and P. Mogensen, "Beyond 5G wireless IRT for industry 4.0: Design principles and spectrum aspects," in *2018 IEEE Globecom Workshops (GC Wkshps)*, IEEE, 2018, pp. 1-6.

[148] *H. Xie, Z. Qin, G. Y. Li, and B.-H. Juang, "Deep learning enabled semantic communication systems," IEEE Transactions on Signal Processing, vol. 69, pp. 2663–2675, 2021.*

[149] *H. Xie and Z. Qin, "A Lite Distributed Semantic Communication System for Internet of Things," IEEE Journal on Selected Areas in Communications, vol. 39, no. 1, pp. 142–153, Jan 2021.*

[150] *N. Farsad, M. Rao, and A. Goldsmith, "Deep Learning for Joint Source-Channel Coding of Text," in 2018 IEEE International Confer- ence on Acoustics, Speech and Signal Processing (ICASSP). IEEE, Apr 2018.*

[151] *Z. Weng and Z. Qin, "Semantic communication systems for speech transmission," IEEE Journal on Selected Areas in Communications, vol. 39, no. 8, pp. 2434–2444, 2021.*

[152] *E. Bourtsoulatze, D. B. Kurka, and D. Gunduz, "Deep Joint Source- channel Coding for Wireless Image Transmission," in ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, May 2019.*

[153] *J. Wang, Y. Duan, X. Tao, M. Xu, and J. Lu, "Semantic perceptual image compression with a laplacian pyramid of convolutional networks," IEEE Transactions on Image Processing, vol. 30, pp. 4225– 4237, 2021.*

[154] *D. Huang, X. Tao, F. Gao, and J. Lu, "Deep learning-based image semantic coding for semantic communications," in 2021 IEEE Global Communications Conference (GLOBECOM), 2021, pp. 1–6.*

[155] *D. B. Kurka and D. G¨und¨uz, "Deepjscc-f: Deep joint source-channel coding of images with feedback," 2020.*

[156] *X. Li, J. Shi, and Z. Chen, "Task-driven semantic coding via reinforce- ment learning," IEEE Transactions on Image Processing, vol. 30, pp. 6307–6320, 2021.*

[157] *T.-Y. Tung and D. G¨und¨uz, "Deepwive: Deep-learning-aided wireless video transmission," 2021.*

[158] *H. Xie, Z. Qin, G. Y. Li, and B.-H. Juang, "Deep learning enabled semantic communication systems," IEEE Transactions on Signal Processing, vol. 69, pp. 2663–2675, 2021.*

[159] *E. Bourtsoulatze, D. B. Kurka, and D. Gunduz, "Deep Joint Source- channel Coding for Wireless Image Transmission," in ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, May 2019.*

[160] *J. Wang, Y. Duan, X. Tao, M. Xu, and J. Lu, "Semantic perceptual image compression with a laplacian pyramid of convolutional networks," IEEE Transactions on Image Processing, vol. 30, pp. 4225– 4237, 2021.*

[161] *D. Huang, X. Tao, F. Gao, and J. Lu, "Deep learning-based image semantic coding for semantic communications," in 2021 IEEE Global Communications Conference (GLOBECOM), 2021, pp. 1–6.*

[162] *S. Iyer, R. Khanai, D. Torse, R. J. Pandya, K. M. Rabie, K. Pai, W. U. Khan, and Z. Fadlullah, "A survey on semantic communications for intelligent wireless networks," Wireless Personal Communications, vol. 129, no. 1, pp. 569–611, 2023.*

[163] M. T. Ribeiro, S. Singh and C. Guestrin, "" Why should i trust you?" Explaining the predictions of any classifier," in *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, 2016, pp. 1135-1144.

[164] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," *Advances in neural information processing systems,* vol. 30, 2017.

[165] A. Shrikumar, P. Greenside and A. Kundaje, "Learning important features through propagating activation differences," in *International conference on machine learning*, 2017, pp. 3145-3153.

[166] M. Palmonari and P. Minervini, "Knowledge graph embeddings and explainable AI," *Knowledge Graphs for Explainable Artificial Intelligence: Foundations, Applications and Challenges,* vol. 47, no. IOS Press Amsterdam, p. 49, 2020.

[167] A. Verma, V. Murali, R. Singh, P. Kohli and S. Chaudhuri, "Programmatically interpretable reinforcement learning," in *International Conference on Machine Learning*, PMLR, 2018, pp. 5045-5054.

[168] T. Shu, C. Xiong and R. Socher, "Hierarchical and interpretable skill acquisition in multi-task reinforcement learning," *arXiv preprint arXiv:1712.07294,* 2017.

[169] G. Liu, O. Schulte, W. Zhu and Q. Li, "Toward interpretable deep reinforcement learning with linear model u-trees," in *Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2018, Dublin, Ireland, September 10--14, 2018, Proceedings, Part II 18*, Springer, 2019, pp. 414-429.

[170] J. Ali-Tolppa, S. Kocsis, B. Schultz, L. Bodrog and M. Kajo, "Self-healing and resilience in future 5G cognitive autonomous networks," in *2018 ITU Kaleidoscope: Machine Learning for a 5G Future (ITU K)*, IEEE, 2018, pp. 1-8.

[171] T. Omar, T. Ketseoglou and I. Naffaa, "A novel self-healing model using precoding \& big-data based approach for 5G networks," *Pervasive and Mobile Computing,* vol. 73, p. 101365, 2021.

[172] O. Onireti, A. Zoha, J. Moysen, A. Imran, L. Giupponi, M. A. Imran and A. Abu-Dayya, "A cell outage management framework for dense heterogeneous networks," *IEEE Transactions on Vehicular Technology,* vol. 65, no. 4, pp. 2097-2113, 2015.

[173] J. Moysen and L. Giupponi, "A reinforcement learning based solution for self-healing in LTE networks," in *2014 IEEE 80th Vehicular Technology Conference (VTC2014-Fall)*, IEEE, 2014, pp. 1-6.

[174] J. Guo, Z. Wang, X. Shi, X. Yang, P. Yu, L. Feng and W. Li, "A deep reinforcement learning based mechanism for cell outage compensation in massive IoT environments," in *2019 15th International Wireless Communications \& Mobile Computing Conference (IWCMC)*, IEEE, 2019, pp. 284-289.

[175] P. Yu, X. Yang, F. Zhou, H. Li, L. Feng, W. Li and X. Qiu, "Deep reinforcement learning aided cell outage compensation framework in 5G cloud radio access networks," *Mobile Networks and Applications,* vol. 25, pp. 1644-1654, 2020.

[176] W. Feng, Y. Teng, Y. Man and M. Song, "Cell outage detection based on improved BP neural network in LTE system," in *11th International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM 2015)*, IET, 2015, pp. 1-5.

[177] H. T. Oğuz, A. Kalaycioğlu, and A. Akbulut, "Femtocell outage detection in multi-tiered networks using LSTM," in *2019 11th International Conference on Electronics, Computers and Artificial Intelligence (ECAI)*, ieee, 2019, pp. 1-5.

[178] P. Yu, F. Zhou, T. Zhang, W. Li, L. Feng and X. Qiu, "Self-organized cell outage detection architecture and approach for 5G H-CRAN," *Wireless Communications and Mobile Computing,* vol. 2018, 2018.

[179] "TALON," [Online]. Available: https://talon-project.eu/

[180] "5G-PICTURE," [Online]. Available: https://www.5g-picture-project.eu/

[181] L. Abeni, A. Balsini and T. Cucinotta, "Container-based real-time scheduling in the linux kernel," *ACM SIGBED Review,* vol. 16, no. 3, pp. 33--38, 2019.

[182] Z. Liao at al., "Distributed Probabilistic Offloading in Edge Computing for 6G-Enabled Massive Internet of Things," *IEEE Internet of Things Journal,* vol. 8, no. 7, pp. 5298-5308, April 2021.

[183] Z. A. Traspadini, M. Giordani and M. Zorzi, "UAV/HAP-Assisted Vehicular Edge Computing in 6G: Where and What to Offload?," in *Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit)*, Grenoble, France., 2022.

[184] K. Jaiswal, A. Dahiya, S. Saxena, V. Singh, A. Singh and A. Kushwaha, "A Novel Computation Offloading Under 6G LEO Satellite-UAV-based IoT," in *3rd International Conference on Computation, Automation and Knowledge Management (ICCAKM)* , Dubai, United Arab Emirates, 2022.

[185] M. S. Z. Thu and E. C. Htoon, "Markov Based Computational Tasks Offloading Decision for Face Detection," in *IEEE 14th International Conference on Intelligent Systems and Knowledge Engineering (ISKE)*, Dalian, China, 2019.

[186] Y. Zhang, X. Dong and Y. Zhao, "Decentralized Computation Offloading over Wireless-Powered Mobile-Edge Computing Networks," in *IEEE International Conference on Artificial Intelligence and Information Systems (ICAIIS)*, Dalian, China, 2020.

[187] K. Zang et. al., "Energy-efficient offloading for mobile edge computing in 5G heterogeneous networks," *IEEE Access,* vol. 4, p. 5896–5907, 2016.

[188] K. Kim, Y. M. Park and C. S. Hong, "Machine Learning Based Edge-Assisted UAV Computation Offloading for Data Analyzing," in *International Conference on Information Networking (ICOIN)*, Barcelona, Spain, 2020.

[189] S. Ajmal, M. B. Muzammil, A. Jamil, S. M. Abbas, U. Iqbal and P. Touseef, "Survey on Cache Schemes in Heterogeneous Networks using 5G Internet of Things," in *3rd International Conference on Future Networks and Distributed Systems (ICFNDS '19)*, Paris, 2019.

[190] S. Mehamel, K. Slimani, S. Bouzefrane and D. M., "Energy-Efficient Hardware Caching Decision Using Fuzzy Logic in Mobile Edge Computing," in *6th Internationa Conference on Future Internet of Things and Cloud Workshops*, Barcelona, 2018.

[191] M. Chen, W. Li, G. Fortino, Y. Hao, L. Hu and I. Humar, "A Dynamic Service Migration Mechanism in Edge Cognitive Computing," *ACM Transactions on Internet Technologies,* vol. 19, no. 2, p. 15, 2019.

[192] S. Ghosh and P. D. Agrawal, "A high performance hierarchical caching framework for mobile edge computing environements," in *IEEE Wireless Communications and Networking Conference (WCNC)*, Nanjing, 2021.

[193] M. Guo and W. Wang, "A review of computing offloading and caching in edge computing based on reinforcement learning," in *4th International Conference on Computing, Networks and Internet of Things (CNIOT '23)*, Xiamen , 2023.

[194] Y. Lie and B. Mao, "On a Novel Content Edge Caching Approach based on Multi-Agent Federated Reinforcement Learning in Internet of Vehicles," in *32nd Wireless and Optical Communications Conference (WOCC)*, Newark, 2023.

[195] T. Zhang, X. Fang, Z. Wang, Y. Liu and A. Nallanathan, "Stochastic Game Based Cooperative Alternating Q-Learning Caching in Dynamic D2D Networks," *IEEE Transactions on Vehicular Technology,* vol. 70, no. 12, pp. 13255-13269, 2021.

[196] F. Jiang, Z. Yuan, C. Sun and J. Wang, “Deep Q-Learning-Based Content Caching With Update Strategy for Fog Radio Access Networks,” *IEEE Access,* vol. 7, pp. 97505-97514, 2019.

[197] European Commission, "Broadcast and Multicast Communication Enablers for the Fifth-Generation of Wireless Systems - 5G-Xcast," [Online]. Available: https://cordis.europa.eu/project/id/761498

[198] European Commission, "CacheMire: Wireless Edge Caching Platform - CacheMire," [Online]. Available: https://cordis.europa.eu/project/id/727682

[199] "Primo-5g – EUK-02-2018-PriMO-5G," [Online]. Available: https://primo-5g.eu/

[200] J. Lelli, C. Scordino, L. Abeni and D. Faggioli, “Deadline scheduling in the Linux kernel,” *Software: Practice and Experience,* vol. 46, no. 6, pp. 821--839, 2016.