Cost-Efficient 5G Non-Public Network Roll-Out: The Affordable5G Approach

L. Sarakis[1], P. Trakadas[1], J. Martrat[2], S. Prior[3], O. Trullols-Cruces[4], E. Coronado[5], M. Centenaro[6], G. Kontopoulos[7], E. Atxutegi[8], P. Gkonis[1], S. Gonzalez-Diaz[2], A. Antonopoulos[4], S. Siddiqui[5], P. Merino[9]

[1]General Department, National and Kapodistrian University of Athens, Psachna, Greece. Emails: {lsarakis, ptrakadas, pgkonis}@uoa.gr
[2]Atos, Spain. Emails: {josep.martrat, sergio.gonzalezd.external}@atos.net
[3]Acceleran, Belgium. Email: simon.pryor@accelleran.com
[4]NearbyComputing S.L., Barcelona, Spain. Emails: {otrullols, aantonopoulos}@nearbycomputing.com
[5]i2CAT Foundation, Barcelona, Spain. Emails: {estefania.coronado, shuaib.siddiqui}@i2cat.net
[6]Athonet S.r.l., Bolzano Vicentino, Italy. Email: marco.centenaro@athonet.com
[7]Eight Bells Ltd., Nicosia, Cyprus. Email: george.kontopoulos@8bellsresearch.com
[8]Nemergent Solutions SL, Spain. Email: eneko.atxutegi@nemergent-solutions.com
[9]University of Malaga, Spain. Email: pedro@lcc.uma.es

Abstract—**The era of 5G broadband wireless networks is inextricably connected with the provision of high data rates to mobile users, as well as bandwidth demanding and low latency applications. While large scale deployments of 5G Public Networks are ongoing, enterprises are interested to deploy their own 5G Non Public Networks (NPNs), customized to better serve their specific use cases. To this end, the goal of this paper is to present an architectural approach for cost-efficient 5G Standalone NPN deployment, leveraging cell densification, disaggregated RAN with open interfaces, edge computing and AI/ML-based network optimization. For this purpose, open solutions, such as O-RAN and MANO frameworks for cloud native micro-service deployments are adopted. Furthermore, research, development and deployment challenges are also discussed.**

Keywords—5G; O-RAN; NFV; Standalone Non-Public Networks; AI/ML-based network optimization

I. Introduction

The requirements of three types of services have become the driving force for innovation towards 5G networks. These types include massive Machine Type Communications (mMTC), supporting millions of Internet of Things (IoT) devices with intermittent activity and transmissions of small data packets, Ultra-Reliable and Low-Latency Communications (URLLC), allowing zero-latency communication with high reliability, and enhanced Mobile Broadband (eMBB), accommodating traffic with high data rates, as well as cell-edge users' connectivity. In the last years the rollout of the first phase of 5G (NSA Non-Standalone, eMBB) has been proven critical in many verticals. However, the real impact of the Standalone (SA) 5G system, including the whole spectrum of services and network slices (URLLC and mMTC) as well as non-public deployments, is still to come in many services and business models.

Businesses and industries are looking forward to 5G Non-Public Networks (NPNs) to get high-level granular views of their operations, service flexibility and spread of deployment possibilities or cost reduction in a specialized IoT market, especially if such 5G network is affordable and easy-to-manage. Although the concept of dedicated mobile NPNs for large enterprises or industries is not a new one, the advent of 5G networks provides plenty of new business opportunities for multiple actors [1]. 5G NPNs are expected to deliver high-speed, low-latency and other benefits, supporting next-generation applications. They also ensure that critical civil functions and business processes have access to high-quality and responsive communications, even when parts of the system fail due to external factors. In practice, ensuring continuity and adaptability of service in critical networks means building secure networks with high availability and reliability together with the supporting cloud-native deployment capabilities.

In this context, a technical overview of 5G NPNs is provided in [2], while spectrum opportunities and design challenges are discussed as well. According to [3], simplification of infrastructure and spectrum will encourage the exponential growth of scenario-specific cellular networks (e.g., private networks, community networks, micro-operators). The study in [4] discusses how 5G NPNs can be deployed across diverse licensed, shared licensed, and unlicensed spectrum bands.

Along with the evolution in the 5G capabilities (codified in 3GPP releases like Rel-15 and Rel-16), there is also an industry-wide change towards software defined and cloud technologies, using Commercial Off-The-Shelf (COTS) compute and networking infrastructure to i) manage costs and expand the supplier eco-system, and ii) enhance openness, competition and spur innovation in the Radio Access Network (RAN) and Core Network (CN).

At the same time, initially driven by 5G Mobile Network Operators (MNOs) cost and affordability concerns, the industry transition towards an open, disaggregated, intelligent, virtualized and highly extensible 5G vRAN architecture has consolidated around the O-RAN Alliance [5] vision of complementing 3GPP 5G standards with a foundation of virtualized network elements, white-box hardware and standardized interfaces that fully embrace the core principles of intelligence and openness. An ecosystem of innovative new products is already emerging that will form the underpinnings of the multi-vendor, interoperable, autonomous RAN that have been envisioned before but is only now becoming a reality for 5G [6], [7]. To accommodate this service and vendor heterogeneity, network resource orchestration, flexible slicing and network optimization facilitated by Artificial Intelligence/Machine Learning (AI/ML) and data analytics are essential elements of a 5G network and have received considerable attention in the context of recent EU 5G PPP projects in Phase 2 [8] (like 5G-ESSENCE and 5GCity) and Phase 3 [9] (like 5G-EVE, 5G-CLARITY, ARIADNE, 5G-COMPLETE and MonB5G).

In view of the above, the goal of this paper is to present an architectural approach for 5G Stand-alone NPNs (SNPNs), i.e., NPNs that are independent of public mobile networks, targeting at cost-efficient network deployment and management. This architecture, which is investigated in the context of the EU 5G PPP Affordable5G project [10], is fully exploitable and open by adopting RAN functions on open interfaces and standard hardware platforms. At the same time, the support of innovative verticals will be achieved through cloud native orchestration and slice management, as well as the migration of Virtual Network Functions (VNFs) across different network nodes. In the proposed system, intelligent management of the network, the infrastructure resources and the services are facilitated by AI/ML and the provision of data analytics.

The rest of this manuscript is organized as follows. Key points for a cost-efficient 5G NPN deployment are provided in Section 2. The proposed architectural approach is described in Section 3, while Section 4 summarizes research, development and deployment challenges for NPNs. Finally, concluding remarks are provided in Section 5.

II. Key Enablers for Cost-efficient 5G NPN Deployment

The effective deployment of 5G networks will require hundreds of thousands of new cell sites, new or upgraded connective nodes and central switches, new software and redesigned mobile devices [11]. Network sharing will remain one of the most significant cost mitigators in the 5G era [12], [13]. The use of neutral hosts could be especially attractive in locations with physical space constraints for multiple networks deployments. Dense small cell deployments (as may be required for 5G New Radio - NR service) can be capital intensive to build out and maintain a backhaul and cell site infrastructure that may consist of 10's to 100's of nodes deployed in a relatively small area such as a downtown region, a venue or enterprise location.

The conceptual enablers with a key role in a full effective, cost-efficient 5G deployment can be analyzed as follows.

Network sharing and service topologies. Network sharing is a key concept, as sharing the network infrastructure and adopting network virtualization can help operators to save significant amounts of capital and operational costs [14]. Network topologies have a direct impact on edge deployments and service performance. In addition, such topology lessens the environmental impact, reduces power consumption, mitigates citizens' concern over radiation and decreases the visual impact of each operator's equipment in the same location.

Neutral hosting with particular opportunity for private networks. Currently, neutral hosts offer mostly infrastructure for public networks, i.e., towers, power, RF front-ends and antennas from existing Wi-Fi and 4G technologies. However, we envisage that neutral hosts will be also involved in the management of 5G equipment (small cells) for private networks. With the introduction of 5G services and the system architecture evolution to NFV/Software Defined Networking (SDN), the cost efficiencies of deploying 5G services may be leveraged by a neutral host service provider to provide tailored and differentiated services blended with services offered by MNOs and to maintain continuity of these services within the coverage area of the neutral host.

Servicing hybrid profile and partitioning edge and cloud, including requirements like latency and capacity. In this context, a major technical challenge includes the cost-latency balance choices when positioning edge computing servers from the network edge/cell site (higher cost and lower latency) to the central office/data center (lower cost and higher latency) [15], [16].

Slicing capabilities and dimensioning for service execution. Another significant technical challenge includes the support and operation of different kind of services with very distinct needs onto the same infrastructure in a cost-efficient way. The goal is to avoid over-dimensioning, that has been the norm in the past for avoiding any kind of congestion. Proper planning, dimensioning and enforcement is needed to make the transition to this new form of service sustainable, starting with an appropriate data collection on resource usage, including most significantly virtual ones.

Automation and close control loops as a means of keeping smart operation. This feature includes the design of algorithms that can autonomously manage all the main phases of the 5G slice and edge resource lifecycle and automate the operations through Artificial Intelligence-driven operations (AIOps), reinforcing high-level policies defined by telecom platform providers and verticals to react or to pro-act to the exposed Key Performance Indicators (KPIs) and events to meet Quality of Service (QoS) and operation requirements, while minimizing the costs [17].

Use of open software platforms. Such platforms covering network functionalities at RAN, Edge, core and management, are essential for a cost-effective and reusable network architecture [18].

III. Proposed Affordable5G Architecture

The proposed approach for an O-RAN-based, 5G SA network for SNPN deployments is shown in Fig. 1. The functionalities of the system are grouped in three layers, which are described in the remainder of this section.

The Infrastructure Layer mainly consists of the Core and Edge/Regional NFV Infrastructures (NFVIs), the cell site platform and Transport Network (TN) segments. The NFVIs, i.e., cloud infrastructures where software is decoupled from hardware, may have different capabilities in terms of compute, storage and network resources, and support for virtualization (Virtual Machines - VMs vs. containers) and hardware acceleration engines. The cell site platform includes proprietary infrastructures, where the software is coupled with the hardware and support the deployment of Physical Network Functions (PNFs). The transport part includes optical transport components along with synchronization delivery elements. Configuration and management of the transport links can be done through an SDN controller.

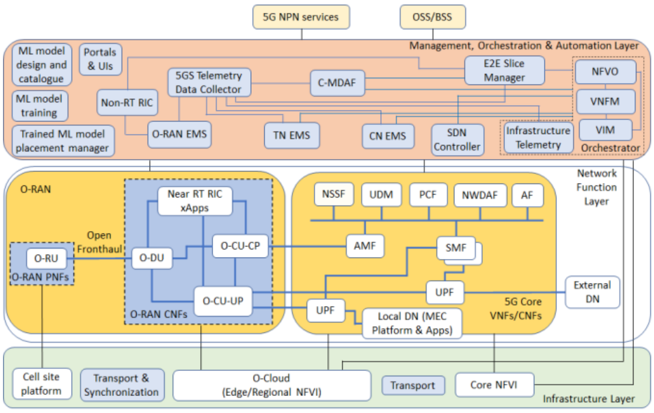

Fig. 1. Proposed system architecture

The Network Function Layer consists of all Network Functions (NFs), related either to the O-RAN or the 5G Core. The O-RAN architecture [19] is based on the decomposition of elements into NFs (either PNFs or Cloud-native NFs – CNFs) and functional splits, and supports separation of the control and user plane functions. These functions include the O-RU (O-RAN Radio Unit) hosting Low-PHY layer and RF processing, the O-DU (O-RAN Distributed Unit), which hosts RLC/MAC/High-PHY layers, the O-CU-CP (O-RAN Central Unit – Control Plane) hosting the RRC and the control plane part of the PDCP protocol, and the O-CU-UP (O-RAN Central Unit – User Plane) that hosts the user plane part of the PDCP protocol and the SDAP protocol. The communication between O-RU and O-DU is provided by the Open Fronthaul interface (Split option 7.2) based on eCPRI. The radio part also includes the O-RAN near Real-Time RAN Intelligent Controller (near-RT RIC), which is a logical function that enables near-real-time control and optimization of O-RAN elements and resources.

The 5G SA core network combines the necessary NFs for the user plane (UPF) and the control plane (AMF, SMF, PCF, UDM, NSSF). The SMF is able to establish multiple PDU sessions that allow the User Equipment (UE) to communicate with multiple UPFs and access services in different Data Networks (DNs), including local DN hosting Multi-access Edge Computing (MEC) applications. The PDU session types include IPv4, IPv6, Ethernet and unstructured, depending on the type of the transmitted information. The 5G Core also includes the Network Data Analytics Function (NWDAF) which is responsible for providing network analysis information upon request from NFs. The role of this function will be discussed in Section III.C.

The main functionalities at the Management, Orchestration and Automation (MOA) Layer are detailed in the next subsections.

A. Orchestration

The NFV Orchestrator (NFVO) performs onboarding and lifecycle management of Network Services (NS), including instantiation, scaling, updating and termination. The VNF Manager (VNFM) is responsible for overseeing the lifecycle management of VNF instances, while the Virtual Infrastructure Manager (VIM) Controls and manages the compute, storage, and network resources [20].

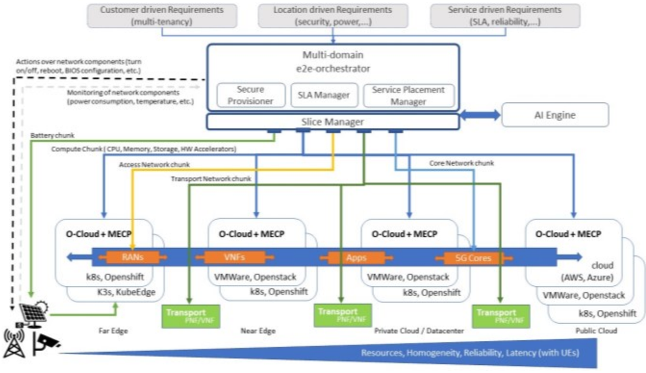

Fig. 2. Orchestrator components

The envisaged orchestrator (Fig. 2) provides a complete end-to-end orchestration solution in real 5G infrastructures, composed of heterogeneous components, i.e., from virtual network functions to MEC resources and hardware devices, spanning across multiple domains with various underlying technologies (i.e., Openstack and K8s). It is responsible for joint network and compute resource orchestration, actuation over network infrastructure/components and CPU pinning for isolation and guaranteed QoS. Furthermore, the orchestrator allows for service migration and pinning to nodes with specific HW accelerators or features (e.g., lower latency). Example of such accelerators include shared vGPU resources between concurrent VMs and low-power consumption Vision Processing Units (VPUs) for inference.

On top of these solutions, advanced AI-enabled algorithms are incorporated to facilitate network automation, minimizing the human intervention towards zero-touch network provisioning. Finally, the proposed orchestration tool is aligned with the O-RAN specifications and compatible with the relevant interface for managing O-Clouds (i.e., O2), allowing the easy integration towards open and fully interoperable mobile networks.

B. Slicing

The slice manager is able to provision end-to-end slices in the compute, network, and access network domains, allowing several tenants to seamlessly manage the required resources, and to deploy services for different verticals within the slices. The slice manager is located at the MOA layer, as it can be seen in Fig. 3, which depicts the high-level view of this component. Within this layer, the main goal of the slice manager is the creation and management of network slice instances. Such instances are composed of three chunks of resources, namely compute, networking and radio.

The slice manager is responsible for the resource partitioning of the slices at three levels, as defined by the NGMN Alliance [21]: infrastructure, network slice, and vertical service. At the infrastructure level, the slice manager receives the slices requirements and, based on those, the radio and computing virtualization infrastructure is created and activated through the interaction with the VIM and the non Real-Time RAN Intelligent Controller (non-RT RIC). This operation reserves the resource chunks and associates them under a specific end-to-end slice ID. At network slice level, the network services required for the communication establishment of the infrastructure are instantiated. In this phase, the core network is activated, and the small cells start radiating. Moreover, the slice is fully activated by enabling the end-to-end connectivity along all elements. At this moment, the UEs have full connectivity. Finally, the vertical service level is optional and allows third-party services to be deployed in the slice through the NFVO and the VNFM. In this specific case, the slice manager offloads the VNFs deployment, monitoring, scaling and migration operations on the slices to the orchestrator.

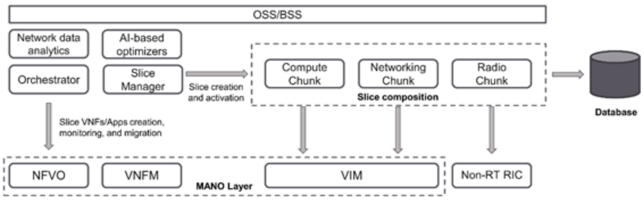

Fig. 3. High-level view of the network slicing system

The slice manager includes a native AI component located at the orchestration and automation layer. This component has as main goal to analyze the information from the 5G System (5GS) Telemetry Data Collector in Fig. 1 and to provide concrete information on the status of the slice resources. Such function will provide periodic information on the status of the infrastructure, the RAN and the 5G CN. The AI extension will build on this data to predict possible scarcities in the resources of the slices in a specific time frame that could prevent the deployment of new services in the slice or even put at risk the performance of the running ones.

C. Telemetry and Data analytics

The 5GS Telemetry Data Collector in Fig. 1 is responsible for the monitoring of NF (both virtual and physical) “application layer” performance metrics (through the interaction with the respective Element Management System – EMS – of the different networks) and the collection of related data from the physical and virtualized infrastructure via the Infrastructure Telemetry component of the orchestrator. In this context, the proposed system supports performance data collection from OpenStack environments (i.e., VMs) and Kubernetes environments (i.e., containers), together with the ones gathered from the different network elements [22].

The data collected at the 5GS Telemetry Data Collector are made available for analysis operations like event correlation, anomaly detection, performance monitoring, metric calculation, trend analysis, and other advanced processes, such as AI/ML algorithms that target at improving specific KPIs of the network. In the proposed system, data analytics are provided by the NWDAF and the Centralized Management Data Analytics Function (C-MDAF).

The NWDAF [23] combines information from different 5G Core NFs, as well as infrastructure telemetry data, and provides analytics and predictions that can be used by 5G Core NFs (like PCF and NSSF). This NF was introduced by the 3GPP, as part of the 5G core architecture, to address the lack of standardized interfaces for data collection for analytics purposes, as well as the delivery of analytics services. The function enables network operators to either implement their own ML-based data analytics methodologies or integrate third-party solutions to their networks. NWDAF incorporates standard interfaces from the service-based architecture to collect data by subscription or request data from other network functions and similar procedures. NWDAF can either provide network analysis data to other NFs (i.e., analytics information) or NFs can request subscription from NWDAF for data delivery (i.e., events subscription) by using a standardized interface.

The C-MDAF, on the other hand, provides Management Data Analytics Services (MDAS) [24] that can be consumed by various entities, like management service producers/consumers for network and service management, NFs (e.g., NWDAF), Self-Organizing Network (SON) functions, network and service optimization tools, service level assurance functions, human operators, and Application Functions (AFs). In the system, predictions and recommendations provided by the C-MDAF can indicatively be used by the slice manager and the orchestrator to address potential system performance degradation through appropriate resource and service reconfigurations.

Both the NWDAF and the C-MDAF can utilize AI/ML for advanced data analytics facilitating intelligent decision making. To support such ML-based data analytics, the proposed system includes facilities for ML model design, training, packaging, and model placement in the appropriate component.

D. AI/ML-based RAN optimization

Current sophisticated 5G systems are too complex to manage, which led service providers to be interested in solutions that allow them to operate their network with minimal interventions. In this respect, intent-based RAN management aims at allowing service providers to specify the connectivity service and prioritize among users and services, based on business intent.

The Non-RT RIC is a new entity introduced by O-RAN Alliance to achieve Intent-based management, through automation and AI/ML tools. Its tasks include configuration management, device management, fault management, performance management and lifecycle management for all network elements. In particular, the non-RT RIC enables non-real-time control and optimization of RAN elements and resources, as well as execution of AI/ML workflow including model training, inference and updates. The non-RT RIC consists of a platform along with a set of microservices (called rApps) and provides policy-based guidance and AI/ML model management to the near-RT RIC through the A1 interface [25].

To give a clearer picture of the role of these RICs in ML-based network optimization, we discuss an example workflow scenario for a Deep Learning (DL) Radio Resource Management (RRM) task (Fig. 4). In a high-level point of view, the workflow may be described in seven generic steps. In the initial phase of the deep learning cycle (step 1), centralized and distributed O-RAN nodes provide data to the data collector entity located in the MOA layer through the O1 interface [26]. Optionally, the collected data may be pre-processed prior to the model training. In step 2, the pre-processed model is forwarded to Non-RT RIC, while in step 3, the model training procedure is hosted in the AI platform (e.g., AcumosAI), upon request. Upon successful completion of the model training, the model is sent back to the Non-RT RIC, thus becoming available for inference purposes (step 4). In step 5, through the A1 interface, the pre-trained model is forwarded to the Near-RT RIC for possible inference hosting. During near-real time control, the model predictions are applied to the O-RAN actors, when the model inference triggers a solution (step 6). Finally, in step 7, the ML-assisted network entities provide feedback metrics to the MOA for possible model update/re-training.

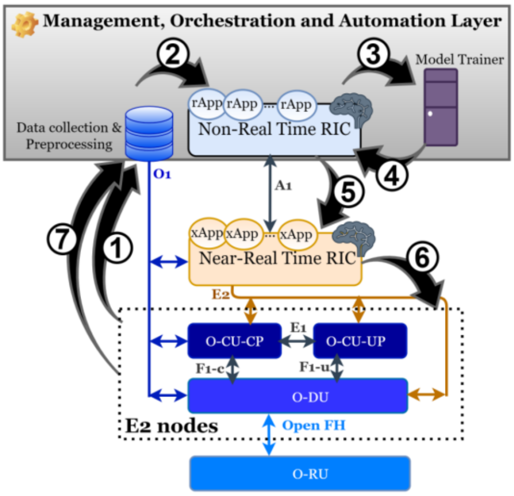

Fig. 4. ML-based workflow executed at the O-RAN

As readily observed in this architecture, the O-RAN intelligence workflow follows a bottom-up approach by hierarchically hosting the computationally expensive procedures (e.g., training) in the MOA layer and the Non-RT RIC, while the near-zero operations (e.g., inference and actions) are performed at the centralized, distributed and radio units of the O-RAN.

IV. Research, Development and Deployment Challenges

The realization of the 5G SNPN discussed in the previous section poses significant challenges that involve not only the development of the disaggregated RAN and the 5G Core but also the efficient management and sharing of available cloud resources, the tailoring of the system to the needs of specific applications and use cases, the AI/ML-based network and service delivery optimization, the support of available hardware acceleration platforms, as well as security. We elaborate on some of these challenges in the following subsections.

A. Support of O-RAN and placement of NFs

The use of standardized interfaces for O-RAN element and O-Cloud management is important to avoid vendor lock-in. The implementation of the O1 interface will allow the non-RT RIC to collect related info from various components of O-RAN using a standardized approach. In addition, related data can be also gathered directly from the Infrastructure Layer via the O2 interface through which the O-Cloud management is performed. Specification of this interface has not been completed by the O-RAN Alliance; however, related guidelines have been provided and should be followed by implementations claiming compliance with the O-RAN.

Regarding the O-RAN and the UPF NFs placement, there is currently an increased interest to analyze the implications of their deployment in the available cloud-native infrastructures (edge, regional and core cloud). The deployment depends on various factors such as the available fronthaul capacity, traffic density, the number of supported RUs and network slices, the need for pooling of centralized O-DU resources, the existence of local DN, the presence of multi-tenant neutral host with multiple operators sharing O-RU etc.

B. Orchestration and network slicing

With respect to orchestration, most of the existing solutions focus on the orchestration of network resources and data centers, without explicitly considering the challenges of the network edge. In such an environment, storage, communication and computational resources are scarcer and, thus, more valuable. In addition, the consideration of network edge makes the network truly heterogeneous, stressing the need for orchestrating different functions (both physical and virtual) and different devices (e.g., sensors, cameras, etc.). Moreover, as more technological domains are incorporated in the network, the possibility of having multiple administrative domains and underlying infrastructures (e.g., VMs or containers) increases significantly, requiring high level orchestrators that orchestrate the individual domain orchestrators.

The 5G SNPN architecture proposed in this work highlights the need for efficient slice management facilitated through an AI/ML-based component that can provide predictions regarding resource utilization impacting the quality of service of established slices. The integration of this simple yet effective AI extension within the 5G network management platform will pave the way for further enhancements such as prediction of events to pre-empt mitigation measures, dynamic spectrum sharing, etc.

C. AI/ML and analytics for network optimization

An interesting direction is the deployment and testing of both DL and Deep Reinforcement Learning (DRL) models towards the resource management of O-RAN, utilizing the learning capabilities of both Non-RT and Near-RT RICs. Such models can be tested in either centralized or distributed manner, with variable modelling scenarios for the agent/learner and/or objective functions. For example, physical layer-based optimization algorithms can be employed to manage the physical resources, such as the user association rules and the power configurations of RUs. Additionally, a MAC-layer (or higher) based optimization method can be an end-to-end network slicer, optimizing not only the RAN slice resources (physical resource block allocation for each slice), but also the virtual resources of the core network (assignment of VNFs to different or shared slices).

In this context, implementation of a complete training-inference-retraining cycle tested and checked within the whole parts of O-RAN architecture is important. Using the proposed system, we aim at a holistic testing of ML models, starting from the model deployment and training, model inference and action taking, and ending at the model evaluation for possible retraining and update.

Regarding the support for network and management data analytics, open-source solutions for the realization of the NWDAF and the C-MDAF, able to gather requested information from NFs across all the layers of the proposed system based on standardized interfaces, would encourage and allow multivendor deployments and facilitate customization to suit individual service needs. In addition, these data analytics functions can be integrated with an AI framework (like AcumosAI); this integration will allow to develop, test, validate and evaluate ML models against specific use case requirements.

D. Hardware acceleration

Introduction of specialized (Application-Specific Instruction set Processor – ASIP-type) hardware accelerators combining i) efficient processing of Deep Neural Network (DNN) workloads, ii) low power consumption, and iii) performance scalability is also challenging for network operations at the edge relying on AI and DNN algorithms (e.g., video processing). Such a hardware accelerator could also be used to host the ML model inference and appropriately be adapted in different O-RAN components (e.g., in the near-RT RIC as hardware-based xApp or in the DU for distributed inference for computationally expensive convolutional neural network-based models).

E. Security

5G NPNs are built as a critical part of enterprise's infrastructure, and security is a driving factor for the ground-breaking services that will take place in the 5G ecosystem. The novel virtualization techniques that enable an efficient and flexible utilization of the different infrastructures add an extra layer of complexity to the mobile network ecosystem. In this context, security must be assessed at all levels, from the physical resources, passing through the virtualization infrastructure up to the network services.

V. CONCLUSIONS

An architectural approach has been presented, encapsulating various scientific and technological challenges surrounding the design, implementation and cost-efficient deployment of 5G NPNs. The analysis identified several key enablers and related challenges for the realization of NPNs, such as software RAN based on open interfaces, network sharing, end-to-end orchestration in infrastructures composed of heterogeneous components, network slicing and AI/ML and analytics-based network automation and optimization.

NPNs enable QoS customization (coverage, throughput, latency, uplink/downlink ratio, KPI values) while addressing requirements for increased availability, reliability and security for many use cases. In this context, the proposed system will be validated in several operational scenarios, including i) industrial use cases with time synchronization constraints (leveraging Time Sensitive Networking – TSN over 5G), ii) mission critical services including voice, video and data, and iii) video analysis at the edge for smart city applications.

ACKNOWLEDGMENT

This work has been partially supported by the Affordable5G project, funded by the European Commission under Grant Agreement H2020-ICT-2020-1, number 957317 through the Horizon 2020 and 5G-PPP programs (www.affordable5g.eu/).

REFERENCES

[1] A. Rostami, "Private 5G Networks for Vertical Industries: Deployment and Operation Models," *2019 IEEE 2nd 5G World Forum* (*5GWF*), Dresden, Germany, 2019, pp. 433-439.

[2] A. Aijaz, "Private 5G: The Future of Industrial Wireless," *accepted for publication in IEEE IEM* (available at: https://arxiv.org/pdf/2006.01820.pdf)

[3] M. N. Patwary, S. Junaid Nawaz, M. A. Rahman, S. K. Sharma, M. M. Rashid and S. J. Barnes, "The Potential Short- and Long-Term Disruptions and Transformative Impacts of 5G and Beyond Wireless Networks: Lessons Learnt From the Development of a 5G Testbed Environment," *in IEEE Access*, vol. 8, pp. 11352-11379, 2020.

[4] G. Brown, "Private 5G Mobile Networks for Industrial IoT," Qualcomm (White Paper), available at: https://www.qualcomm.com/media/documents/files/private-5g-networks-for-industrial-iot.pdf

[5] O-RAN ALLIANCE, https://www.o-ran.org

[6] S. E. El Ayoubi (Ed.), "5G Innovations for New Business Opportunities," 5G-PPP, March 2017.

[7] G. Caruso, F. Nucci, O. P. Gordo, S. Rizou, J. Magen, G. Agapiou, P. Trakadas, "Embedding 5G solutions enabling new business scenarios in Media and Entertainment Industry," 2019, *IEEE 2nd 5G World Forum (5GWF)*, pp. 460-464.

[8] 5G PPP Phase 2 Projects, https://5g-ppp.eu/5g-ppp-phase-2-projects

[9] 5G PPP Phase 3 Projects, https://5g-ppp.eu/5g-ppp-phase-3-projects

[10] The Affordable5G project: High-tech and affordable 5G network roll-out to every corner, 5G PPP, https://www.affordable5g.eu

[11] J. F. Valenzuela-Valdés, Á. Palomares, J. C. González-Macías, A. Valenzuela-Valdés, P. Padilla and F. Luna-Valero, "On the Ultra-Dense Small Cell Deployment for 5G Networks," *2018 IEEE 5G World Forum* (*5GWF*), Silicon Valley, CA, 2018, pp. 369-372.

[12] A. Antonopoulos, "Bankruptcy Problem in Network Sharing: Fundamentals, Applications and Challenges," IEEE Wireless Communications, vol. 27, no. 4, pp. 81-87, August 2020.

[13] N. Nomikos et al., "A UAV-based moving 5G RAN for massive connectivity of mobile users and IoT devices," *Vehicular Communications*, vol. 25, 2020, 100250.

[14] W. S. H. M. W. Ahmad et al., "5G Technology: Towards Dynamic Spectrum Sharing Using Cognitive Radio Networks," *IEEE Access*, vol. 8, pp. 14460-14488, 2020.

[15] X. Wang, Y. Han, V. C. M. Leung, D. Niyato, X. Yan and X. Chen, "Convergence of Edge Computing and Deep Learning: A Comprehensive Survey," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 2, pp. 869-904, Second quarter 2020.

[16] P. Trakadas et al., "Hybrid clouds for data-intensive, 5G-enabled IoT applications: An overview, key issues and relevant architecture," *Sensors*, vol. 19, issue 16, 2019.

[17] M. E. Morocho-Cayamcela, H. Lee and W. Lim, "Machine Learning for 5G/B5G Mobile and Wireless Communications: Potential, Limitations, and Future Directions," *in IEEE Access*, vol. 7, pp. 137184-137206, 2019.

[18] F. Kaltenberger, A. P. Silva, A. Gosain, L. Wang and T. T. Nguyen, "OpenAirInterface: Democratizing innovation in the 5G Era," *Computer Networks* (*Elsevier*), vol. 176, July 2020, 107284.

[19] O-RAN Alliance, "O-RAN Architecture Description," O-RAN.WG1.O-RAN-Architecture-Description-v02.00, October 2020.

[20] P. Trakadas et al., "Comparison of management and orchestration solutions for the 5G era," *Journal of Sensor and Actuator Networks*, vol. 9, issue 1, 2020.

[21] NGMN Alliance, "Description of Network Slicing Concept," 2016. (available at https://www.ngmn.org/wp-content/uploads/160113_NGMN_Network_Slicing_v1_0.pdf)

[22] P. Trakadas et al., "Scalable monitoring for multiple virtualized infrastructures for 5G services," *International Symposium on Advances in Software Defined Networking and Network Functions Virtualization, (*SoftNetworking), 2018.

[23] 3GPP, "5G System; Network Data Analytics Services; Stage 3," 3GPP TS 29.520, V15.3.0, April 2019.

[24] 3GPP, "Study on enhancement of Management Data Analytics (MDA) (Release 17)", 3GPP TR 28.809, V2.0.0, March 2021.

[25] O-RAN Alliance, "A1 interface: Application Protocol," O-RAN.WG2.A1AP-v02.00, July 2020.

[26] O-RAN Alliance, "O-RAN Operations and Maintenance Interface Specification," O-RAN.WG1.O1-Interface.0-v03.00, March 2020.