AI-powered Infrastructures for Intelligent and Privacy-aware Beyond-5G Systems

Leonardo Militano, Anastasios Zafeiropoulos, Roberto Bruschi, Andy Edmonds, Eleni Fotopoulou,
Chiara Lombardo, Symeon Papavassiliou, and Thomas M. Bohnert

***Abstract*—In this paper, a vision for beyond-5G systems is proposed where automation, intelligence and data privacy in cloud-native infrastructures are in focus. Exploiting the convergence of cloud technologies at the edge and mobile communication networks, a set of architectural and technological solutions is discussed that will play a fundamental role on the path from 5G towards future sixth-generation systems. Currently, a strong need is felt in the telecommunication world for greater automation to meet the extreme requirements expected for future 6G applications. In this regard, Artificial Intelligence (AI) is gaining high momentum as one of the central enabling technologies for beyond-5G networks. Reinforcement Learning (RL) and Federated Learning (FL) are here proposed as technologies to enhance network automation and enable privacy-aware applications. Blockchain is proposed as a solution for non-repudiation and trustworthiness in the AI pipelines. These technologies are brought together in a comprehensive cloud-native architectural vision to fill the gap between current 5G systems and AI-powered secure systems of the future.**

***Index Terms*—6G systems, Federated Learning, Artificial Intelligence, Reinforcement Learning, Blockchain, Automated Orchestration, Edge Computing, Cloud-native.**

I. Introduction

FIFTH generation (5G) networks are becoming the core enabler for the information society of 2020, but researchers are already in the process of drafting ideas about what the intelligent information society of 2030 will be with the sixth generation (6G) of wireless networks. Many 5G concepts are currently being applied and delivered, whereas new concepts and technologies will soon arise to differentiate 6G networks. NFV (Network Functions Virtualization) capabilities, together with SDN (Software Defined Networking) mechanisms and edge computing are among the key 5G technologies. They are crucial in enabling and supporting the deployment and orchestration of a wide range of vertical applications with very heterogeneous but extremely challenging performance and operating requirements. On the other hand, the huge amount of data collected by sensors, embedded in all sorts of terminals, machines, and things, has to be networked with low latency fixed-radio connections, elaborated in the Cloud and Edge-Fog Computing facilities, to eventually result into a variety of ICT (Information and Communication Technology) services. Therefore, 5G systems are taking on the characteristics of a powerful networking-computing-storage infrastructure with functions partly distributed and partly centralized, supporting pervasive connections (wired and wireless) characterized by high capacity and very low latency (a few milliseconds).

If, on one side, technologies first introduced in 5G networks (e.g., softwarization, cloudization, virtualization, slicing) [1] will still play an important role in autonomous networks of the future, *intelligentization* and *connected intelligence* are the recurring keywords to distinguish the next generation networks from the past [2] [3]. It is, therefore, not surprising that pervasive Artificial Intelligence (AI) is deemed to be among the central enabling technologies for beyond 5G networks [2] [4] [5]. We are observing how the telecoms world already provides an innovative baseline infrastructure, which spans from the core out to the radio base stations and, ever increasingly, to edge devices. However, as scale increases, so does management and the overhead of that. That complexity is already present in the horizontal aspects of the network management, but when vertical end-user-owned applications are deployed, this complexity explodes adding advanced orchestration needs (e.g., reactive scaling decisions, self-* functionalities). Therefore, *the need for automation and dynamic orchestration of services and resources is strongly felt*. Moreover, extensions of current and beyond-5G NFV-based systems, will have to *embed secure and trustworthy* solutions into the management of core network services and end-user applications to allow the operation on user data *without loss of privacy and trust*.

Considering these aspects, the key research question we address in this paper is: *"How to enable intelligent and dynamic management of infrastructures, network services and applications through data processing over virtualized remote and local resources, to enable greater trust in providers, protect the end-users privacy, and achieve high performance?"*

II. Societal & Technical Challenges on the Road to 6G Systems

We can identify some societal, business and technical challenges to be addressed by next-generation systems.

- **Societal and Business Considerations.** Today the *exchange of data over networks* is a common occurrence. This data intrinsically has value, and when transformed

L. Militano, A. Edmonds and T.M. Bohnert are with the Zurich University of Applied Sciences (ZHAW), Switzerland (e-mail: milt@zhaw.ch, edmo@zhaw.ch, bohe@zhaw.ch).

R. Bruschi is with the University of Genoa, DITEN Department, Italy (e-mail: roberto.bruschi@unige.it

C. Lombardo is with the CNIT S2N National Laboratory, Italy (e-mail: chiara.lombardo@cnit.it)

A. Zafeiropoulos, E. Fotopoulou and S. Papavassiliou are with the National Technical University of Athens (NTUA), Greece (e-mail: tzafeir@cn.ntua.gr, efotopoulou@netmode.ntua.gr, papavass@mail.ntua.gr)

through processing, provides highly useful insights to the operation of a system. Nonetheless, with its value comes the use of customer data in a way unknown to customers. This is typically more evident in business-to-customer interactions. Protection of the end-users, be they developers or citizens, is something that needs to be considered more carefully in upcoming 6G architectures. Transparency of data and decision-making processes is also important in some vertical industries, making necessary the development of AI mechanisms that can be self-explanatory.

- **Convergence of cloud-native and edge technologies.** To reap the benefits of decentralization, the convergence of cloud-native infrastructures and mobile networks must be enhanced in all components of compute, storage and networking. Connectivity should not only be reliable, but also able to adapt to changing applications' and networks' requirements. While 5G systems already moved a step in this direction with several functionalities for Edge Computing integration, such as the User Plane Function and the Session Management Function defined in the 3GPP 5G Service-Based Architecture, actual implementations are far from maturity. At the same time, computation and storage at the edge of the network are facing an exponential data growth which greatly increases the requirements for faster data processing. Therefore, novel cloud-native storage and computation solutions are seeing the light to guarantee low power and efficient computation (e.g., computational storage, in-memory storage/computation, and multi-tier hybrid storage).
- **Automated orchestration and management.** In 5G systems, customized network instances, called network slices, are assigned to vertical stakeholders to interconnect their applications hosted at the edge and have the ability to manage them even if their components are spread over multiple domains. Currently, slicing often relies on the same orchestration mechanisms developed in the past which do not provide the agility required by upcoming applications. In this respect, AI is the most promising technique to dynamically provide the required feedback for real-time adaptations in automated orchestration. Decision-making for orchestration actions has to be assisted or fully delegated to AI mechanisms. Reinforcement Learning can be considered as an indicative and promising technology for this purpose.
- **Security and privacy.** With multiple tenants deploying their applications over an infrastructure potentially owned by multiple providers, security and privacy must be guaranteed by-design. Federated Learning (FL) and Blockchain are two technologies that can help answering to this need, as FL systems have privacy-by-design and blockchain technology can enable non-repudiation of model contributions. With FL, raw data will be stored and operated upon locally on devices, whereas only model updates and refinements are sent to an external aggregation node for further processing and update of the global model. In all of this, no data is ever shared with the central nodes and it never leaves the data holder's possession, whereas the blockchain non-repudiation feature will track the participating nodes and the metadata/data sources.

III. 6G Enabling Technologies: Federated Learning, Reinforcement Learning and Blockchain

A. *Federated Learning for Distributed Model Training*

FL is a specific machine learning (ML) technique able to put privacy front-and-centre [6]. In contrast to traditional centralized ML techniques where all data samples are uploaded to one centralized server for model updates computation, FL trains an algorithm across multiple decentralized devices holding local data samples. This prevents the privacy-violating transport of user/application data to untrusted external nodes "bringing the code to the data, instead of the data to the code" [7]. So-called FL server nodes will collect the model updates (the metadata) produced by the end-devices, merge them and produce an updated global model. A further positive effect FL introduces is that raw data is not sent over bandwidth-constrained networks to central nodes of the network infrastructure improving latency, energy consumption and environmental impact. Moreover, the network infrastructure nodes are not being overloaded with computation, storage and communication demanding tasks. How the metadata is going to be exploited depends on the domain and the vertical application. Nonetheless, the possibility to share the metadata among similar domains and contexts paves the way for multi-application and multi-providers collaborations.

Referring to the FL lifecycle described in [6], three main phases are considered: Selection, Configuration and Reporting. Fig. 1 reports a representation of one FL round.

Selection Phase: an FL pipeline starts with devices (worker nodes) that meet some eligibility criteria (e.g., enough computational power and energy levels), checking-in to the server and announcing that they are ready to run an FL computation task for a given FL application. Among the many available devices during a certain time window, the server selects a subset of devices based on certain objective functions to work on a specific FL task for a given round (Step 1).

Configuration phase: The devices stay connected to the server for the duration of the FL round and get instructions from the server node about what computation to run and how to execute it. The server node sends to each participant the current global model parameters and any other necessary state (Step 2).

Reporting phase: Each participant performs a local computation based on the global state and its local dataset (Step 3) and sends a model update back to the server (Step 4). The server aggregates these updates into its global state (Step 5) and reports the devices when to reconnect. The process repeats for the next FL rounds (Step 6).

The described lifecycle is an ideal case where no network or device failures occur. When the reporting from some worker nodes is not possible, e.g., due to network failure/poor connectivity or mobility, the FL is interrupted. To solve this issue, *we propose the adoption of D2D communications [8] to*

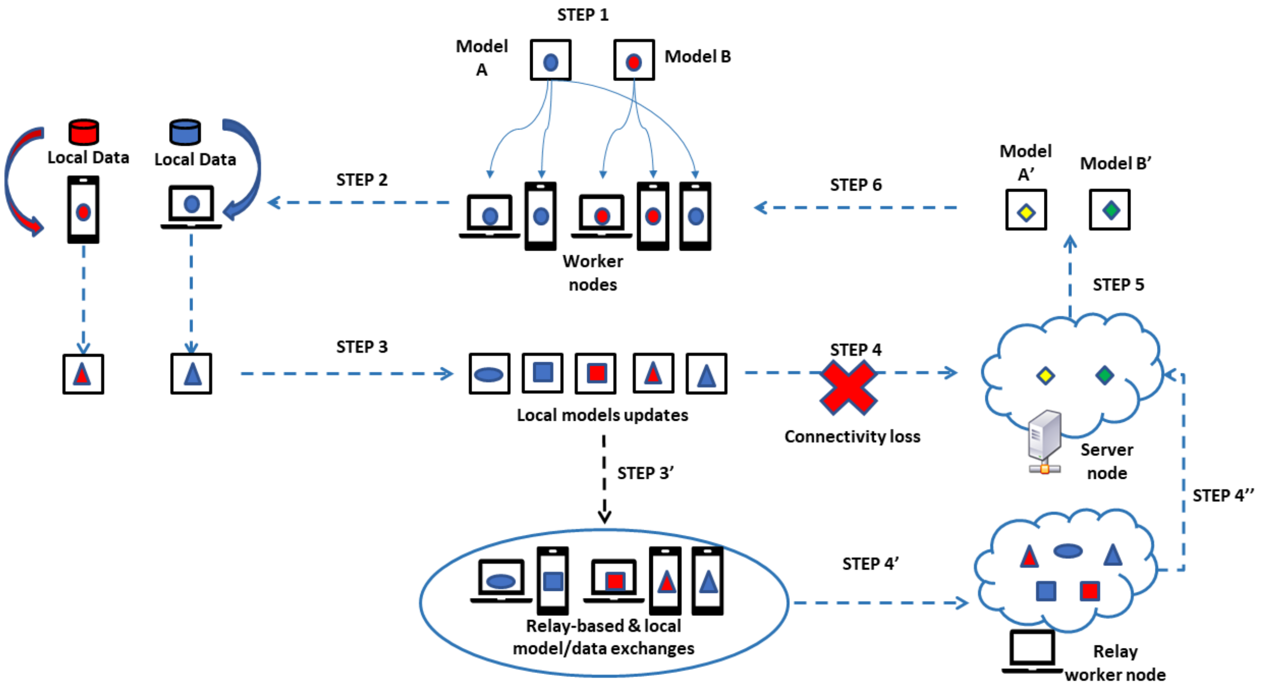

Fig. 1. D2D-aided Federated Learning lifecycle for AI-model updates at the edge.

enable a relay node to act as a temporary collector of model updates for FL worker nodes. The relay could be the node with the available (or best) connectivity to the server node, or the one with the largest computation/battery/storage resources. Such a D2D-based FL can also be leveraged to further improve the FL protocol itself, exploiting data locality information to detect any outliers or anomalies in the trained partial models caused by biased samples.

B. Reinforcement Learning for Automated Orchestration

Reinforcement Learning (RL) is a generic framework for representing and solving control tasks, where decisions must be made or some behaviour must be enacted [9]. In RL, the learning algorithm decides which actions to take for a control task, based on the definition of an ultimate goal to achieve. If the goal is accomplished, a reward is provided, while the objective of the algorithm is to maximize the long-term reward. Input data, bundled in the form of states, is provided by the environment following a dynamic process. Modeling of a control task in RL is realised as a Markov decision process (MDP), where each decision does not require knowledge of the prior states and actions.

RL can boost the development of intelligent orchestration mechanisms and introduce automation and self-learning characteristics in 6G network management, tackling aspects related to optimal deployment, network slice and elasticity management of network functions and services. The blending of RL and FL technologies is also promising, since FL can provide valuable data for training and evaluation of ML pipelines applied for network management purposes, increasing their accuracy and capacity to efficiently tackle orchestration aspects (see Fig. 2).

Currently, NFV orchestrators (NFVOs) are supporting functionalities for network slices lifecycle management, optimal on-boarding of Virtual Network Functions (VNFs) and network services (NSs) and horizontal/vertical scaling actions. These features enable better usage of the available resources, provision of guarantees for conformance with Service Level Agreements, cost and energy reduction. Nonetheless, *we are still far away from the inclusion of automation in NFV orchestration decision making* and the reduction of the configuration overhead posed to network administrators. For instance, in case of policy-driven elasticity management (e.g., scaling of VMs), mainly rule-based management systems are used, posing overhead for rules declaration to the network administrator. Furthermore, dynamicity in the deployment environment (e.g., deprovision of resources, changes in routing schemes) cannot be easily managed, leading in many cases to far-from-optimal decision making and non-consideration of corner cases (e.g., placing a VNF at a cluster that faces connectivity issues or network attacks).

To address such inefficiencies, *we propose RL mechanisms to support ML pipelines within existing orchestration platforms* and take advantage of the plethora of monitored data. This data is provided by monitoring frameworks of NFVOs and include resources' usage metrics of containers, VMs, clusters and VNF-specific metrics. The great advantage of RL is its capacity to change its behaviour when the environment changes. But this adaptability comes at a price. RL agents not only need a lot of data but also a wide set of experiences to improve their accuracy and avoid models' over-fitting. RL agents have

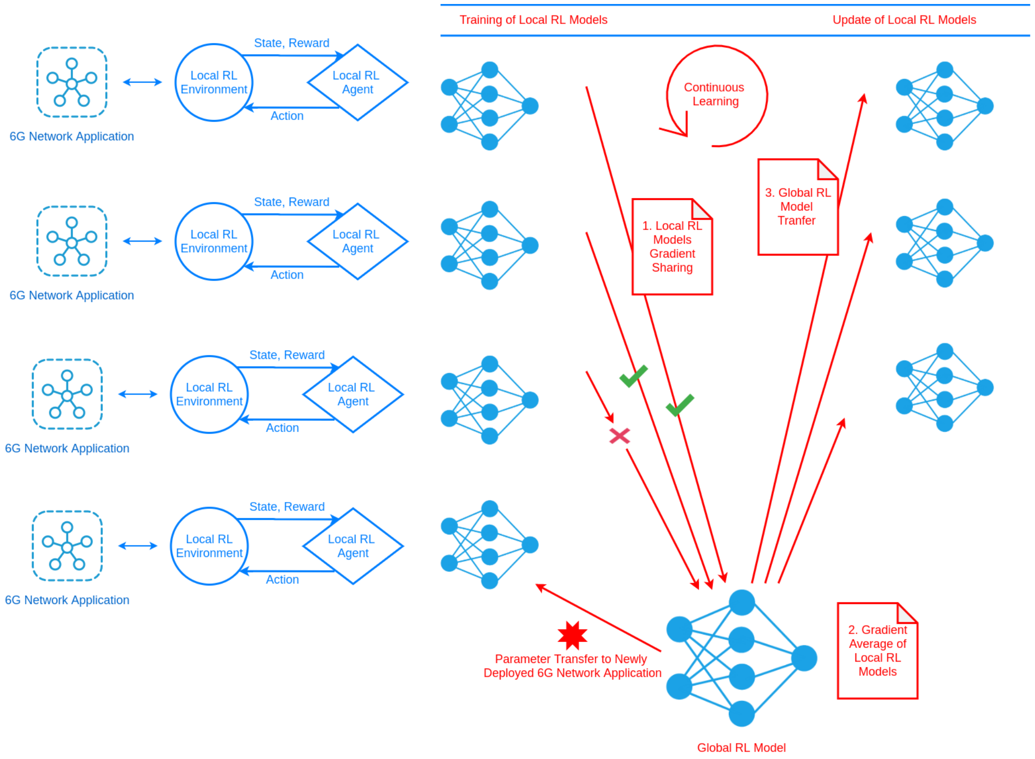
Training of Local RL Models
Update of Local RL Models
6G Network Application
Local RL Environment
Local RL Agent
State, Reward
Action
Continuous Learning
1. Local RL Models Gradient Sharing
2. Gradient Average of Local RL Models
3. Global RL Model Tranfer
Parameter Transfer to Newly Deployed 6G Network Application
Global RL Model

Fig. 2. Reinforcement Learning Model Continuous Updates.

to be optimally trained under a set of diverse deployment scenarios and operational conditions (e.g., different workload characteristics) to provide models that can be generic enough.

To tackle this challenge, we believe that the exploitation of FL pipelines can boost the performance and reliability of RL mechanisms. Aggregating data collected by various local RL agents and updating constantly a global RL model may lead to high efficiency in including automation in orchestration tasks. For instance, in case of elasticity efficiency management, we can produce an RL model per type of VNF, considering their categorization based on resources consumption trends. Such RL models can be trained locally and update the relevant global RL model, leading to automated elasticity efficiency policies based on the type of the VNF in an application-agnostic way. Both the local and global models should be trained online. In this way, data is continually entering the system and is incorporated into the global model through continuous updates. RL models can be updated during runtime, being able to reflect changes in the monitored environment.

Such a process, in the form of a training pipeline, is depicted in Fig. 2 where FL is exploited for collaboratively building RL models. Based on gradients values provided by local RL agents, a global RL model is maintained through the calculation of average gradients' values and made available back to the local RL agents. Based on the latest version of the global model, each RL agent interacts with the local RL environment and provides further updates to the global RL model. Upon the deployment of a new 6G network application, the available model parameters are provided to the local agent to optimally manage its deployment and orchestration.

Among the current challenges, we see a need for RL environments creation, fed by monitoring mechanisms of existing NFVOs. Having a common set of observable metrics will help the community to build intelligent agents that can be trained to opt for sophisticated actions. There is also a need for defining a set of rewards based on how to promote or penalize the possible orchestration actions on behalf of the agents,fine as there is a great lack of literature on how RL agents are incentivized to choose between the different actions they support.

C. Blockchain for Trustworthiness and Non-Repudiation

When it comes to identifying who has access, extracts, and refines data, and what this is going to be used for, concerns soon arise for several stakeholders. First and foremost, data-holders are concerned about their private data being stored and unknowingly exploited in remote cloud-based systems. Therefore, to truly leverage new advanced solutions society-wide, guarantees and assurances need to be given to data-holders such that their data will not be transparently exploited by data-processor services. *We believe that Blockchain is a valid solution for networking infrastructures to securely manage data from FL models [11]*. Being a blockchain resistant to data modification, its implementation will guarantee *authenticity, integrity and non-repudiation* of the information flowing through the infrastructures. A further advantage is that data will never be lost as transactions between parties are recorded efficiently and in a verifiable and permanent way. The participants of a blockchain network will be able to verify and audit transactions independently and relatively inexpensively since data will be available for checks and verification at any time. The decision on which network node is going to be part of the blockchain will have to take into consideration the energy consumption and computational power requirements. This observation would therefore exclude end-user devices, but instead consider more appropriate an edge node in the network infrastructure.

Blockchain can also be adopted for the implementation of rewarding systems based on tokens, to incentivize users to offer their computational power for FL pipelines based on some rewarding model. A tokenization distributed application on the blockchain nodes will generate and assign tokens (prizes) to the most committed devices in the FL system. The earned tokens can be used by the winners to get, for instance, more computational or network resources, triggering a sort of virtuous circle among the participants of the network.

IV. A Vision for AI-powered Privacy-aware 6G systems

In introducing our architectural vision for AI-powered privacy-aware 6G systems of the future, a cautious view on current activities from Standards Development Organizations such as ITU and European Telecommunications Standards Institute (ETSI) on one side, and the evolving requirements of future vertical applications is presented.

A. Standardization activities

NFV MANO [10] is the ETSI-defined framework for the management and orchestration of all resources in a virtualized data center including compute, networking, storage, and virtual machine resources. Its main focus is to allow flexible and easy on-boarding of network components. It is composed mainly of three functional blocks, i.e., the NFVO, the VNF Manager and the Virtualized Infrastructure Manager (VIM). Open Source MANO is an ETSI-hosted initiative to develop an Open Source NFV Management and Orchestration (MANO) software stack aligned with ETSI NFV.

The **ETSI ENI** (Experiential Networked Intelligence Industry Specification Group) [12] defines a Cognitive Network Management architecture that adopts AI and context-aware policies to adjust offered services based on changes in user needs, environmental conditions and business goals. It integrates well with the 5G networks where automated services are operated, with optimized slice management and resource orchestration. Its current focus is on improving the operator experience, using closed-loop AI mechanisms based on context-aware policies for actionable decisions in network management.

The **ETSI ZSM** (Zero-touch network and Service Management Industry Specification Group) [13] aims to provide a framework that enables zero-touch automated network and service management in a multi-vendor environment. The framework is applied on top of specifications made available by NFV, MEC and ENI working groups and aims to incorporate existing and future solutions in a common automation framework, and provide an integration framework towards full end-to-end network service automation.

A closely related standard is the **ITU FG ML5G** (Focus Group on Machine Learning for Future Networks including 5G) [14]. The Focus Group drafted ten technical specifications for machine learning (ML) for future networks, including interfaces, network architectures, protocols, algorithms and data formats. A comprehensive set of (architectural) requirements are identified leading to specific architecture constructs needed.

B. 6G Applications Requirements

6G systems will continue to build on its antecedent technology where the original use cases defined in 5G systems (Enhanced Mobile Broadband -eEMBB, Massive Machine Type Communications -mMTC, and Ultra-Reliable and Low Latency Communications -URLLC) may be combined into a new generation of applications with unprecedented requirements. These will include extreme high speed and high capacity communications, extreme coverage extension, extreme low power consumption and cost reduction, extreme low latency, extreme reliable communication, extreme massive connectivity and sensing.

The architecture proposed in this paper is designed to meet the aforementioned needs. A possible 6G application scenario is that of large-scale enterprises/manufacturers, in which heterogeneous and sensitive data (e.g., images, sensing data) are collected by UEs, IoT sensors, drones. To identify potential problems in the industrial production or working processes, huge amounts of raw data are collected locally and then sent, stored, and processed by centralised/remote cloud services. The proposal we make in this paper aims at keeping the data on the devices within the company administrative domain sharing only the metadata. The computational power is moved from a centralized to a distributed architecture reaching the edge of the network thanks to an intelligent orchestration of the distributed edge/cloud resources. Cloud-native computation and storage technologies involving the edge of the network, will allow for data exchange, latency, and energy consumption reduction at different levels of the infrastructure. Automated orchestration mechanisms will be introduced for supporting re-configuration and scaling actions, based on the guidance of RL mechanisms. Raw data will be processed locally, and only the resulting metadata will be then provided centrally. FL algorithms will reduce latency, improve efficiency and reduce costs. Since the updated global model is then redistributed to all the users of the model, secure access to the data is of paramount importance. For this challenge, blockchain solutions are proposed to guarantee the needed non-repudiation and auditability properties.

C. An AI-powered Privacy-aware Cloud-native Architecture

In Fig. 3 we report our proposal for an architecture integrating the enabling technologies discussed in this paper. The overall objective is to deliver end-to-end trustworthy and privacy-aware AI-powered infrastructures with advanced automation features, leveraging the convergence of cloud-native technologies and 5G networks.

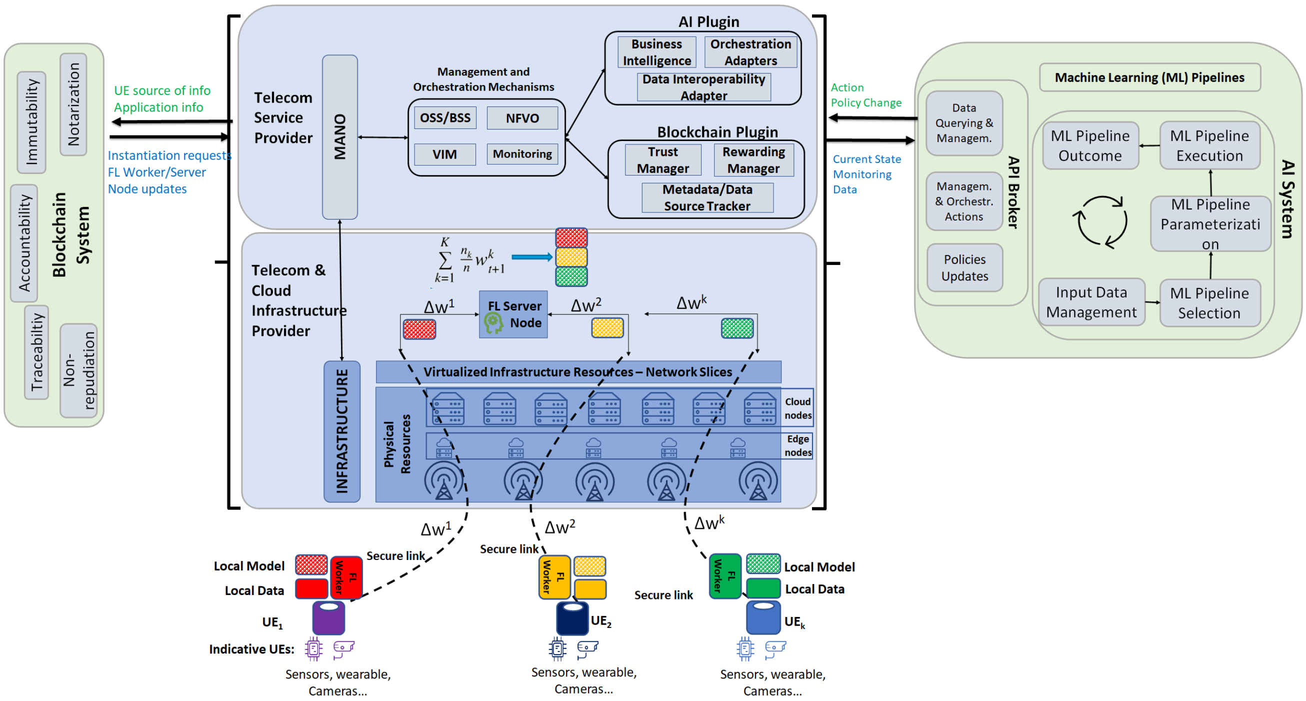

Fig. 3. Proposed Architecture for AI powered and secure 6G systems.

Network operators will benefit of a standards-compliant system that enables pervasive intelligence in their network to increase zero-touch management capabilities to reduce the management burden of ever increasing workload deployments. The infrastructure nodes should be at the edge of the network, as close as possible to the data itself to minimize latency, energy consumption and costs. This is where operators of cloud-like (also NFV-based) infrastructure stand to have an excellent advantage. They own the “last mile”, they are “closest” to end-users, and they will provide the infrastructure of the future.

Application providers will benefit of a solution where the right resources can be acquired and placed close to the customer’s site. End-users efficiently take advantage of Telcos’ distributed infrastructures with dynamic cloud-native infrastructures ranging from the core to the edge of the network. Once deployed, the application can exploit the FL capability whereby the application’s data remains with the end-user and of the RL capability for automated orchestration.

Four main building blocks are identified:

- **NFV MANO**: This is primarily devoted to the deployment of network slices management mechanisms, including the lifecycle management of network services based on NFV principles and the dynamic management of end-to-end network and compute resources for serving the vertical application needs. It includes the deployment, management and orchestration of software elements on the underlying infrastructure through a number of components, namely the Operation Support System (OSS), the Business Support System (BSS), the NFVO, the VIM and the monitoring component. It interacts with the Infrastructure layer to orchestrate the physical and virtual resources, with the AI system, through an AI plugin, for receiving inputs from the ML pipelines, and with the Blockchain system, through a Blockchain plugin, for integrating privacy and trustworthiness into the system.
- **Operator Infrastructure**: This includes the set of physical and virtual resources that the operator must manage and govern. This ranges from server-class physical machines, upon which virtual machines and containers are deployed, storage pools to allow persistence to physical and software controllable network routers that provide the features related to network slicing, all the way to radio heads.
- **AI System**: This is responsible for providing AI capabilities to the MANO. It provides an open and ETSI standards’ compliant interface (e.g., based on specifications of the ENI API-Broker) to ML/FL/RL capabilities. These capabilities are primarily used by the NFVO for supporting AI-assisted orchestration mechanisms and by the OSS/BSS for supporting the extraction of business intelligence analytics.
- **Blockchain system**: This is responsible for providing blockchain-based secure and trustworthy features to both network services and vertical applications. This will integrate trust, non-repudiation and tokenization of rewarding for contributing to AI model training. In case of misbehaving or corrupted devices producing wrong or biased data, the overall FL algorithm can be affected. Here the blockchain non-repudiation feature will be of high importance to track the worker node and the metadata/data source. This is especially important in multi-tenant, multi-site, and even multi-organization scenarios, where malicious or also criminal actions could occur.

In the proposed architecture, the AI plugin is responsible

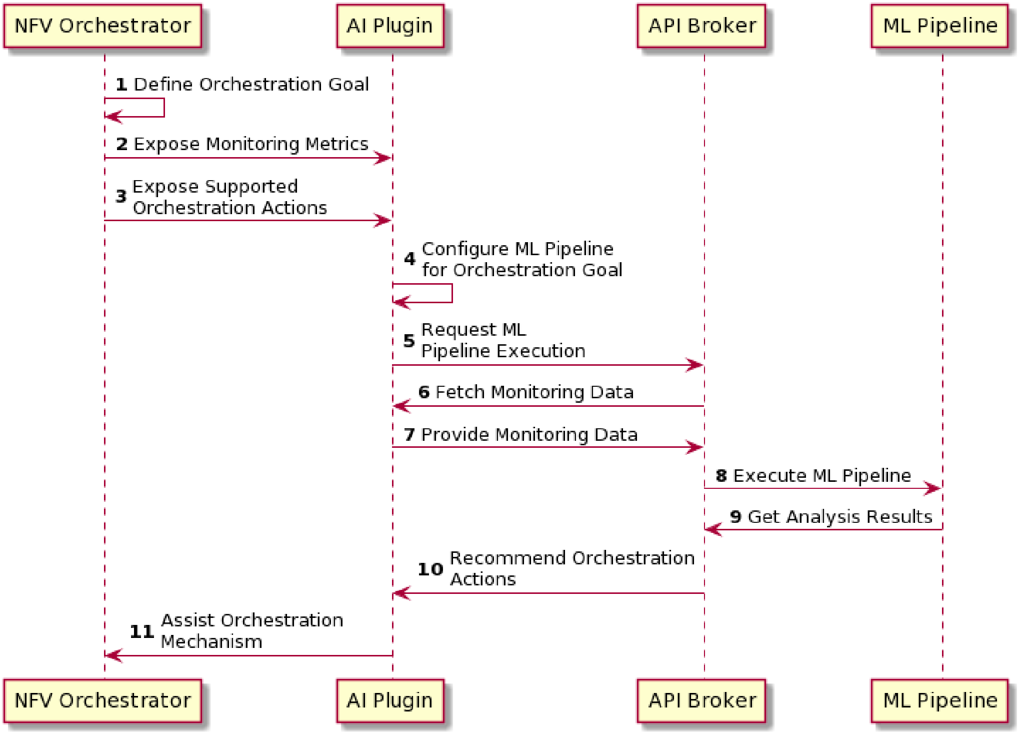
NFV Orchestrator
AI Plugin
API Broker
ML Pipeline
1 Define Orchestration Goal
2 Expose Monitoring Metrics
3 Expose Supported Orchestration Actions
4 Configure ML Pipeline for Orchestration Goal
5 Request ML Pipeline Execution
6 Fetch Monitoring Data
7 Provide Monitoring Data
8 Execute ML Pipeline
9 Get Analysis Results
10 Recommend Orchestration Actions
11 Assist Orchestration Mechanism
NFV Orchestrator
AI Plugin
API Broker
ML Pipeline

Fig. 4. Indicative workflow for a ML Pipeline execution to assist orchestration mechanisms.

to support the injection of AI in decision-making processes of the NFVO and the OSS/BSS, based on the guidance of the AI System. It includes a set of adapters for decision-making with regards to orchestration and business intelligence aspects, and a data interoperability adapter for managing the data provided by the AI System. The AI plugin interacts with the AI System through well-specified application programming interfaces (APIs), as provided by the API Broker (Fig. 4).

Upon the definition of a high-level orchestration goal by a MANO component, AI-assisted orchestration can be realised taking advantage of the execution of a ML pipeline in the AI System. To do so, the AI Plugin is responsible to proceed to proper configuration of the ML pipeline, considering the available monitoring metrics and orchestration actions by the NFVO. Following, data querying and management APIs are used to provide the set of required data for analysis to the API Broker (e.g., time-series data from the Monitoring Engine of the NFVO). The ML pipeline is executed and the provided results are provided back to the AI Plugin in the form of recommendations for undertaking orchestration actions (e.g., scaling out a VNF). Such recommendations are translated to specific actions targeted to the deployed MANO system, aiming to achieve the declared goal. Evaluation of the achieved efficiency is taking place, leading to continuous training and improvement of the applied algorithms. The specified interfaces can be implemented based on existing NFVO implementations (e.g., OSM) and are compliant with the specifications provided by the ETSI ENI ISG. It should be also noted that, in the AI System, ease onboarding and execution of ML pipelines can be supported, including FL and RL pipelines.

Similarly, the Blockchain plugin will support the integration of Blockchain solutions for data/metadata source tracking and trustworthiness management at the MANO layer. The Blockchain plugin will interact with the Blockchain system through open APIs to trigger Blockchain creations, send/receive updates on nodes involved in FL pipelines, manage the tokenization and rewarding of nodes and users, collect information about trustworthiness of users and applications. Decisions on which type of Blockchain will be taken by the OSS, based on the application constraints, to guarantee the wished authenticity, integrity, and non-repudiation features. The BSS instead, will trigger the initiation of tokenization distributed applications with the scope to generate and assign tokens (prizes) to the most committed devices in the FL system.

V. Conclusion

In this paper we discussed research directions and technological trends to address system and application requirements for future 6G systems. With an eye on the current challenges and standardization activities, we discussed how the convergence of cloud-native technologies and mobile communication networks can be exploited to enable highly flexible and dynamic infrastructures ranging from the cloud to the edge of the network. The proposed architectural vision gives, among others, answers to the need for intelligent and automated orchestration, privacy, security and trustworthiness. Infrastructures interfaces, the cloud-native technologies to be adopted and interactions with existing management and orchestration systems have been discussed. RL capabilities that significantly enhance current state-of-the-art solutions for automated orchestration of resources have been analyzed. The growing concerns about end-users data privacy found an answer through the adoption of FL at the edge of the network. Whereas, non-repudiation of AI training contribution and trustworthiness of the involved devices was obtained through a systematic integration of Blockchain technologies.

We believe that the innovative ideas and the solutions discussed in this paper will form an important base of discussion and a starting point for future research activities in industry and academia towards the definition of future 6G systems.

References

[1] Afolabi, I., Taleb, T., Samdanis, K., Ksentini, A., Flinck, H. (2018). Network slicing and softwarization: A survey on principles, enabling technologies, and solutions. IEEE Communications Surveys & Tutorials, 20(3), 2429-2453.

[2] Letaief, K. B., Chen, W., Shi, Y., Zhang, J., Zhang, Y. J. A. (2019). The roadmap to 6G: AI empowered wireless networks. IEEE Communications Magazine, 57(8), 84-90.

[3] Zhang, Z., et al., (2019). 6G wireless networks: Vision, requirements, architecture, and key technologies. IEEE Vehicular Technology Magazine, 14(3), 28-41.

[4] Tariq, F., Khandaker, M. R., Wong, K. K., Imran, M. A., Bennis, M., Debbah, M. (2020). A speculative study on 6G. IEEE Wireless Communications, 27(4), 118-125.

[5] Zhou, Z., Chen, X., Li, E., Zeng, L., Luo, K., Zhang, J. (2019). Edge intelligence: Paving the last mile of artificial intelligence with edge computing. Proceedings of the IEEE, 107(8), 1738-1762.

[6] Bonawitz, K., et al., (2019). Towards federated learning at scale: System design. arXiv preprint arXiv:1902.01046.

[7] Wang, S., Tuor, T., Salonidis, T., Leung, K. K., Makaya, C., He, T., Chan, K. (2019). Adaptive federated learning in resource constrained edge computing systems. IEEE Journal on Selected Areas in Communications, 37(6), 1205-1221.

[8] Militano, L., Araniti, G., Condoluci, M., Farris, I., Iera, A. (2015). Device-to-device communications for 5G internet of things. EAI Endorsed Trans. Internet Things, 1(1), 1-15.

[9] Roig, J. S. P., Gutierrez-Estevez, D. M., Gündüz, D. (2019). Management and Orchestration of Virtual Network Functions via Deep Reinforcement Learning. IEEE Journal on Selected Areas in Communications, 38(2), 304-317, doi: 10.1109/JSAC.2019.2959263.

[10] Network Functions Virtualisation (NFV); Management and Orchestration, ETSI Std. GS NFV-MAN 001 V1.1.1, Dec. 2014.

[11] Kim, H., Park, J., Bennis, M., Kim, S. L. (2019). Blockchained on-device federated learning. IEEE Communications Letters, 24(6), 1279-1283.

[12] Experiential Networked Intelligence (ENI); system architecture", ETSI GS ENI 005 v1.1.1, 2019, http://www.etsi.org/technologies/experiential-networked-intelligence Accessed on October 2020.

[13] Zero touch network and Service Management (ZSM), ETSI ISG, https://www.etsi.org/technologies/zero-touch-network-service-management Accessed on October 2020.

[14] FG ML5G - Unified architecture for machine learning in 5G and future networks, http://www.itu.int/pub/T-FG-ML5G-2019 Accessed on October 2020.

Leonardo Militano is a senior researcher ZHAW School of Engineering in Switzerland where he leads the cloud storage initiative. He received his Ph.D in Telecommunications Engineering in 2010 from the University of Reggio Calabria, Italy. Before joining ZHAW he was an Assistant Professor at the Mediterranea University in Italy, with interests in wireless and mobile networking. His works have been published in more than 70 papers in international journals, conferences and book chapters in the information technology field.

Anastasios Zafeiropoulos (tzafeir@cn.ntua.gr) received the Dipl.-Ing. and Ph.D. degrees from the School of Electrical and Computer Engineering, National Technical University of Athens. Currently he is a PostDoc researcher at the NETMODE Lab at NTUA. He has co-authored over 50 publications in high-level international journals and conferences.

Roberto Bruschi is Associate Professor of Telecommunication Networks at the University of Genoa, DITEN Department. Roberto took part in the activities of many national and European projects. He has co-authored over 130 papers in international journals, book chapters and international conference proceedings, and he was the recipient of two Best Paper Awards, one Runner-Up Best Paper, and one Best Demo Awards. He is a Life Senior Member of IEEE.

Andy Edmonds is a senior researcher at ZHAW School of Engineering in Switzerland. There he leads themes of Infrastructure as a Service and compute. His research interests include distributed and system architectures, virtualization, service-oriented architectures, and cloud computing. He worked in industrial and academic positions in Siemens, Infineon, Intel and the Distributed Systems Group in Trinity College Dublin. He has a research master's degree in distributed systems from Trinity College Dublin.

Eleni Fotopoulou (efotopoulou@netmode.ntua.gr) is a PhD candidate at NETMODE Lab, NTUA Greece, with expertise on the design and development of intelligent orchestration and analytics solutions in the 5G, Cloud Computing, IoT and energy domains. She has more than 15 publications in peer-reviewed international journals and conferences.

Chiara Lombardo received her Ph.D. in Electronics, Informatics, Robotics and Telecommunications Engineering at the University of Genoa in 2014. Chiara has worked as a postdoc at the University of Genoa, DITEN Department for six years and currently works with the CNIT S2N National Laboratory. She has co-authored over 20 papers in international journals, book chapters and international conference proceedings. Her current research interests cover NFV, edge computing and 5G networks.

Symeon Papavassiliou (papavass@mail.ntua.gr) is a professor in the School of ECE at National Technical University of Athens. He has an established record of publications in the field of next generation network modeling, optimization, and management, with more than 300 technical journal and conference papers.

Thomas M. Bohnert (bohe@zhaw.ch) is a Professor with the Zürich University of Applied Sciences. His interests are focused on enabling ICT infrastructures, ranging across mobile/cloud computing, service-oriented infrastructure, and carrier-grade service delivery (Telco + IT). He was with SAP Research, SIEMENS Corporate Technology.