A New Agent-Based Intelligent Network Architecture

(Article begins on next page)

A New Agent-Based Intelligent Network Architecture

Sisay Tadesse Arzo, Domenico Scotece, Riccardo Bassoli, Fabrizio Granelli,
Luca Foschini and Frank H.P. Fitzek

Abstract—The advent of 5G and the design of its architecture has become possible because of the previous individual scientific works and standardization efforts on cloud computing and network softwarization. Software-defined Networking and Network Function Virtualization started separately to find their convolution into 5G network architecture. Then, the ongoing design of the future beyond 5G (B5G) and 6G network architecture cannot overlook the pivotal inputs of different independent standardization efforts about autonomic networking, service-based communication systems, and multi-access edge computing. This article provides the design and the characteristics of an agent-based, softwarized, and intelligent architecture, which coherently condenses and merges the independent proposed architectural works by different standardization working groups and bodies. This novel work is a helpful means for the design and standardization process of the future 5G and 6G network architecture.

I. INTRODUCTION

The International Telecommunication Union (ITU) Telecommunication Standardization Sector (ITU-T) Focus Group on IMT-2020 concluded its pre-standardization activities in December 2016. The core reform that 5G introduced in communication was the idea of virtualization, or more

S. T. Arzo and F. Granelli are with the Department of Information Engineering and Computer Science (DISI), University of Trento, Trento, Italy. (e-mail: {sisay.arzo,fabrizio.granelli}@unitn.it).

D. Scotece and L. Foschini are with Department of Information Engineering and Computer Science (DISI) at University of Bologna, Bologna, Italy. (e-mail: {domenico.scotece,luca.foschini}@unibo.it).

R. Bassoli and F. H.P. Fitzek are with the Deutsche Telekom Chair of Communication Networks, Institute of Communication Technology, Faculty of Electrical and Computer Engineering, Technische Universität Dresden, Dresden, Germany.

F. H.P. Fitzek is also with Centre for Tactile Internet with Human-in-the-Loop (CeTI), Cluster of Excellence, Dresden, Germany. (e-mail: {riccardo.bassoli,frank.fitzek}@tu-dresden.de).

specifically, network softwarization. This implied a tremendous paradigm shift from the previous store-and-forward to the current-future compute-and-forward. This has made computing as important as communication in future communication networks [1]. However, the design of 5G architecture and core characteristics has leveraged the previous experience, obtained during the successful development of cloud computing and network virtualization instances such as software-defined networking (SDN) – led by the no-profit consortium Open Network Foundation (ONF) – and network function virtualization (NFV) – started by the industry and standardized by the European Telecommunications Standards Institute (ETSI).

Recently, ITU has started writing a report about the future technology trends towards 2030 and beyond, which is going to be released in June 2022. This report will mainly provide the very general vision for future 6G communication networks. In 2021, 6G research and design started in the EU, like the EU flagship Hexa-X (hexa-x.eu) and the German research hub 6G-life (6g-life.de). Especially, the Hexa-X project has been publishing the main guidelines for 6G characteristics, use cases, key performance indicators (KPIs), and architecture [2], [3]. However, the last few years of research and standardization efforts in in-network intelligence and network softwarization have been providing the mature input for the architectural aspects of the new generation, which is still an open challenge in [2], [3]. Thus, this article provides an architectural design and guidelines for future 5G, beyond 5G, and 6G networks by leveraging the various standardization works by ETSI and 3GPP, considering the current popular trends on microservice- and agent-based intelligent communication systems. In this vision, all softwarized network functions become atomic entities, playing as autonomous, intelligent, and collaborative agents, in line with

the current trends shown in [2], [3]. To the best of the authors' knowledge, this is the first architectural attempt to unify the four existing individual architectures, microservices-based SDN control plane, ETSI SDN-NFV Management and Orchestration (MANO), ETSI Generic Autonomic Networking Architecture (GANA), and 3rd Generation Partnership Project (3GPP)-ETSI Mobile Edge Computing (MEC), a unique and homogeneous framework for 5G, beyond 5G, and 6G.

II. MOTIVATION AND BACKGROUND

This section provides an overview of the main existing independent architectures.

A. Microservices-based SDN Controller

The centralized nature of the SDN control plane leads to several issues on latency, single point of failure, and overloading [4]. These issues have been partially overcome by using the distributed SDN controller approach. However, the SDN controller is a monolithic system that results inefficient in replicating the SDN system. Moreover, monolithic SDN deployment does not allow dynamic management of SDN components and/or functionalities to (de-)activate according to the scenario. To alleviate this, the work in [5] and [6] decomposed the SDN controller and implemented it as a set of microservices components that can run in a distributed fashion with flexibility in recomposition. In particular, the SDN controller is designed as a composition of logical sub-functions, i.e. microservices. They can share network service load and create a robust system against failures.

The work in [6] presented µONOS architectural framework which is proposed by Open Network Operating System (ONOS) [6]. It is expected to be the next-generation SDN architecture. However, µONOS has been specialized mainly for cloud datacenters scenarios. It is limited to certain technologies and is not all 5G compliant. In particular, µONOS uses gRPC protocols family for microservices intercommunication which is not the standard declared in the 3GPP white-paper v15 [7]. 3GPP specifies REST APIs as the standard-de-facto for services intercommunication. In this regard, we have implemented a microservices-based SDN controller deployment that complies with the 5G standard, see Figure 1 (top left) [5]. Our implementation is

based on Ryu SDN Framework and a first release is available to the community at the link: https://gitlab.com/dscotece/ryu_sdn_decomposition/.

B. ETSI SDN-NFV MANO

ETSI released it's ETSI SDN-NFV MANO architecture [8], to provide a unified architecture effectively combining SDN and NFV characteristics. Figure 1 (top right) depicts its structure. This architecture consists of three main entities, called the Network Management System (NMS), the Network Function Virtualization Infrastructure (NFVI), and the Operation/Business Support Scheme (OSS/BSS). The first manages the virtual network, the second set the resources (hardware or software) that are used to run and to connect virtual network functions, and the third sets the applications used by service providers to provide network services. Each one of these layers has an interface to the MANO entity, which hosts the virtual infrastructure managers (VIMs), that aim at controlling the NFVI resources. Next, the virtual network function manager (VNFM) configures and manages the life cycle of virtual network functions in its network domain. Finally, there is the orchestrator for the NFVI who manages the resources across different VIMs, and subsequently the life cycle of network services. Finally, the virtualization layer groups all the element management entities with their respective virtual network functions (VNFs).

C. ETSI GANA

1) Generic Autonomic Network Architecture: ETSI also unveiled a standard reference architecture as an implementation guide for GANA architectural framework for network automation [9]. The main goal of the GANA reference model is to prescribe the design and operational principles for Decision Elements (DEs) as the drivers for cognitive, self-managing, and self-adaptive network behaviors. ETSI is performing several GANA instantiations onto various target standardized reference network architectures. This is to enable autonomic algorithms to be integrated into the design and implementation of DEs. The integration also aimed at standardizing the autonomic network architectures at four levels as depicted in Figure 1 (bottom left).

Network Level Decision Element: this level is designed to operate the outer closed control loops
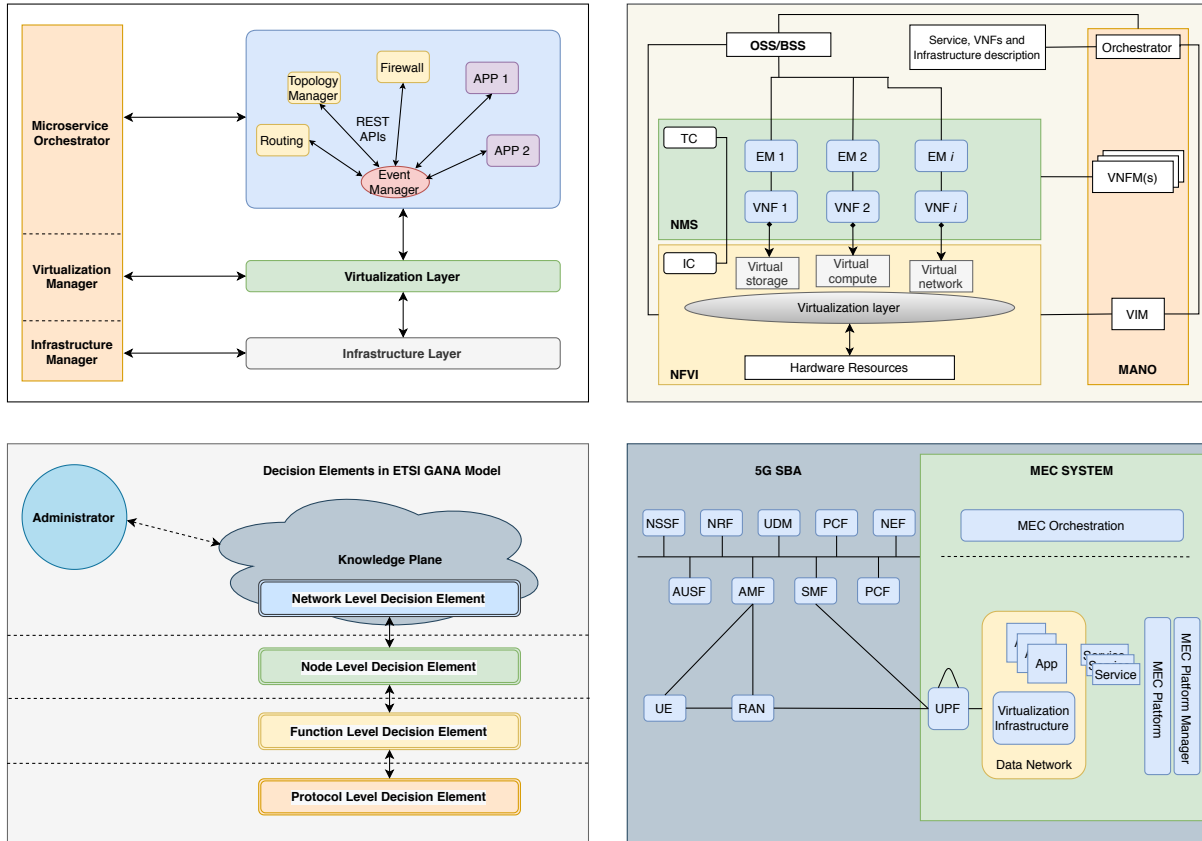
Fig. 1. Architectures of microservices-based SDN Controller decomposition (top left), ETSI SDN-NFV MANO (top right), ETSI GANA (bottom left), and 3GPP-ETSI MEC (bottom right).

on the basis of network wide views or state as input to the DEs' algorithms and logics for autonomic management. **Node Level Decision Element**: this level is in charge of controlling the behavior of the Network Element (NE) as a whole, and also managing the orchestration and policing of the Function Level Decision Elements. GANA Function Level specifies the following four decision elements: security management, fault management, auto configuration and discovery, resilience, and survivability. **Function Level Decision Element**: represents a group of protocols and mechanisms that abstract, as atomic units, networking or a management/control function. The GANA model defines the following six Function Level Decision Elements: routing management; forwarding management; Quality of Service management; mobility management; monitoring management; service and application management. **Protocol Level Decision Element**: this is the lowest level decision element in the system. This kind of element is protocols or other fundamental mechanisms that may exhibit

intrinsic control-loops or decision element logic and associated DE, as is the case for some protocols such as Open Shortest Path First (OSPF). **GANA Knowledge Plane (GANA KP)**: enables advanced management and control intelligence at the Element Management (EM), Network Management (NM), and Operation and Support System (OSS).

A close look at the GANA model reveals the close alignment with the multi-agent-based approach proposed in [10], which defines the network functions as services that can be deployed in a virtualized/containerized environment.

2) Multi-Agent System (MAS): is a sub-branch of distributed artificial intelligence, where it has multiple interacting intelligent agents performing a given task collaboratively and autonomously. MAS includes different attributes such as architecture, communication, coordination strategies, decision making, and learning abilities. The work in [10] identified MAS as a competing candidate for defining atomic and autonomous DEs in the GANA model. DEs could be network function units that could be used as a building block in any automated system.

Network function atomization is the mechanism that defines the smallest possible network function units in a service-oriented system. MAS is comparable to microservices but with more autonomy and proactive capability [10].

D. 3GPP-ETSI MEC Architecture

The basic idea of MEC is to provide capabilities closer to the end-users to overcome mobile difficulties. This promotes a new three-layer device-edge-cloud hierarchical architecture, which is recognized as very promising for several application domains [11]. With the advent of 5G networks, the MEC is one of the key technologies for supporting ultra-reliable and low-latency communications.

The 5G system architecture specified by 3GPP is designed to fit a wide range of use cases including IoT networks management. In particular, the 3GPP defines the Service-Based Architecture (SBA) for the 5G core network, whereby the control plane functionality and common data repositories are delivered by a set of interconnected Network Functions (NFs), each with authorization to access each other's services [12]. The SBA framework provides the necessary functionality to authenticate the consumer and to authorize its service requests, as well as flexible procedures to efficiently expose and consume services. Moreover, ETSI MEC defines an API framework aligned with the SBA framework to provide a set of functionalities and services. The network functions and the services are registered in a Network Resource Function (NRF), while MEC applications are registered in the MEC platform. 5G Network Exposure Function (NEF) acts as a centralized point for service exposure. The Network Slice Selection Function (NSSF) is the function that selects the user slice and allocates the necessary Access Management Functions (AMF). The Unified Data Management (UDM) function generates the 3GPP AKA authentication credentials, while the User Plane Function (UPF) is the configurable data plane. The resulting integrated architecture described in the white paper [12] is presented in Figure 1 (bottom right).

III. PROPOSED ARCHITECTURE

This section discusses the unified architecture, providing conceptual analysis.

A. Proposed Unified Architecture

Our architecture combines four different standards namely: SDN, ETSI NFV, ETSI GANA, and ETSI MEC, for proposing a virtualized network intelligent agents instead of VNFs or microservice (in case of SDN decomposition). Figure 2 shows our unified architecture. While the infrastructure and virtualization layers are directly derived from NFV architecture, the MANO layer is modified to have internal components as intelligent orchestration agents. The application layer of NFV is significantly modified and contains the four layers of the ETSI GANA model. Therefore, the ETSI GANA divides the application layer of NFV into four layers that we further propose the ETSI GANA decision units to be intelligent agents. Hence, the application layer becomes a composition of protocol-level agents, function-level agents, node-level agents, and network-level agents. The event distribution agents enable the event handling at each layer of the GANA model. Note that for event management several strategies are available including centralized, distributed, and hybrid [13].

The application layer depicts the decomposed controller, NMS, and MEC functions on the left and right sides, respectively. The functions are built as autonomous and atomic agents introducing in-network intelligence in the internal architecture of the agents. The recomposition of these agents would create the required controller, MEC system, and other NMS, depending on the requirements in a given environment. The recomposed system becomes a multi-agent-based intelligent system. Network functions in 5G, controller, and NMS and 5G/6G functions should be designed as autonomous agents incorporating appropriate AI/ML as an integral part of that function. The collection of such functions would create a fully autonomous network system. In other words, functions become intelligent agents where intelligence is introduced in the internal architecture of the agent, designing the agents using cognitive components such as AI/ML algorithms.

The interface between the infrastructure layer, virtualization layer, and the application layer is similar to NFV with little modification; however, the orchestration layer is based on intelligent agents. Therefore, we recommended interfaces to be open and adaptive based on the specific applications. This

could be achieved by equipping agents with a Programmable Protocol Stack (PPS), which represents the implementation of a software-based environment of network protocols and layers [14]. The interface between the agents is mainly through restful API. However, we suggested equipping the agent with a dynamic protocol stack.

B. Agent Internal Architecture

We propose agent-designed principles and guidelines to enable intelligence and full autonomy in performing a given function incorporating cognitive components. Figure 3 depicts the general agent design and architectural guideline. An agent is composed by an *INPUT*, which is an incoming request; *FACTS*, is the knowledge database of the agent; *COGNITION (REASONING) UNIT* gives agents the reasoning capability; *PLANNING STRATEGY* organizes the steps/procedures for the action to be taken to satisfy the requested service; *VALIDATION* is the unit that verifies the action plan for consistency, and an *OUTPUT/ACTION* is the final decision (result/action) to be taken by the agent. An agent also has an interface to communicate with other agents or microservices via the PPS. Future generation networks should be equipped with effective and efficient protocols. Agents in the application layer of the proposed architecture are various types that will use combinations of multiple protocols for communication for various kinds of applications including real-time applications. A PPS is a software-based stack that inputs information from the higher logical layers and configures various parameters at each layer, repeating the procedure at each network device. Parameters are updated dynamically according to network changes and service requirements.

C. Network Function Atomization

Network functions can be divided into small functions allowing maximum recomposition freedom. Here we redefine these functions as atomic units having full autonomic capability to perform a given function such as security, QoS monitoring, AMF, SMF, etc. Our architecture re-defines the internal components of a decomposed monolithic system to be intelligent agents, designing the agents using the cognitive element as the brain of the agents to make an autonomous decision on a given function. The agents are fully intelligent, atomic, and autonomous

decision units that can be flexibly recomposed to create a fully autonomic network system. Intelligent refers to the capability of the agent in performing its function, adapting to the dynamic demands. The level of intelligence and capability is heavily reliant on the design of a particular agent. For example, a network traffic classifier and predictor agents can be designed using ML model [15] as a cognitive component of the agent and it can classify and predict given incoming service traffic in order to allocate network resources accordingly depending on the traffic class and proactive allocation and scheduling agents' decision. This reduces resource and energy consumption while reducing service processing latency using proactive resource allocation and scheduling. Again we are assuming task execution agents can optimize their performance depending on the current traffic and system state. Depending on the implemented ML model as a cognitive component of the agent, the performance of the agent varies in terms of accuracy, latency, etc [15]. Similarly, an autonomous SMF agent, designed using an appropriate ML, can dynamically create, update, and remove protocol data unit sessions while managing session context with the UPF agent. Agent atomicity refers to the smallest functionality, identity, and acting territory of a particular function/agent. Autonomy defines the agent's self-reliance in delivering the given task without external intervention. However, the agent can communicate with other agents when it requires information or if completing the task requires the support of other agents. In general, the recomposition of agents would create the overall autonomous system. The recomposition could be agent chaining creation using an orchestration agent. This by itself require intelligence that should be incorporated when designing an intelligent orchestration agent.

D. Multi-Agent Based System Recomposition

As a guideline in designing the agent-based system, we show a decomposed SDN controller [5], designing the components as autonomic agents, and recomposing them to create the intelligent controller system. The main principle to retain in decomposing an SDN controller is that the network information and state should be synchronized and self-consistent. This allows an independent implementation and component reuse. We showed in [5]
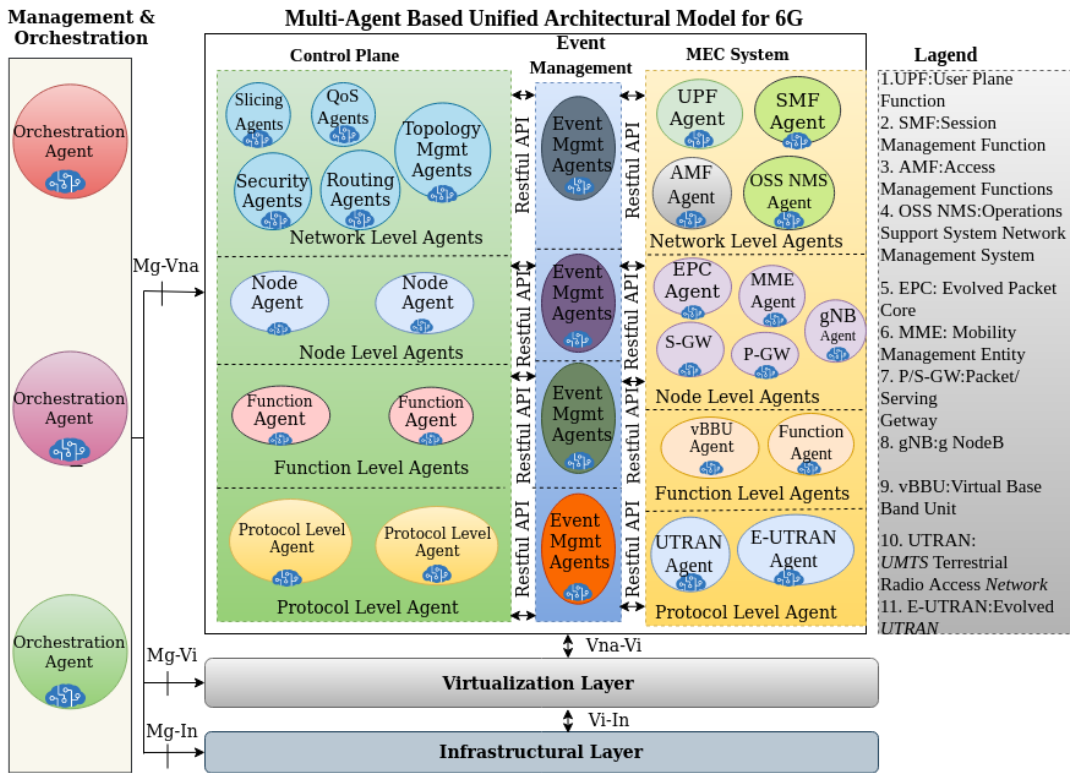
Fig. 2. Microservices and Multi-Agent Based Unified Architectural Model for ETSI GANA, ETSI MEC, ETSI NFV, and SDN
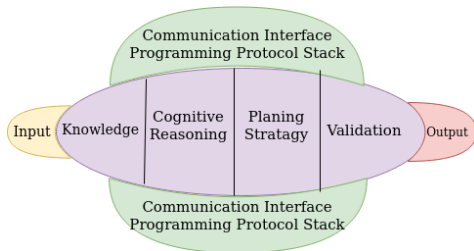

Fig. 3. Agent Internal Architecture

a possible implementation of decomposed SDN Controller using microservice, however, now we re-define these components with intelligent and atomic agents. These agents can be executed on arbitrary computing platforms on distributed and virtualized resources such as virtual machines/containers in edge/cloud environments using intelligent orchestration agents. The loosely-composed system can be viewed as a black box, defined by its externally observable behavior, emulating the original monolithic SDN controller with added intelligence. Agents' orchestration can be done using intelligent orchestrator agents. This creates an agent chain to intelligently perform what the legacy controller does. The recomposition can be done with flexible agent composition as per the system requirement in time and space. In particular, the controller could be deployed with maximum recomposition freedom using only the necessary functions/agents while performing the functions autonomously.

IV. CONCLUSION

This paper presented a unified architecture combining SDN, NFV, MEC, and GANA architectural and conceptual models for future networks, such as beyond 5G, and 6G. The main principles addressed by the architecture are the introduction of in-network intelligence using intelligent agents, decomposing monolithic network systems. Moreover, it also shows the organization of agents to produce an overall automated system. Due to the maximum recomposition freedom and in-network intelligence at the smallest service units of a network system, the architecture is expected to address future network demands in terms of autonomy, flexibility, heterogeneity, reliability, and latency. Since there are few agent based service design one of future challenges this architecture is developing such functions.

REFERENCES

[1] F. H. P. Fitzek, F. Granelli, and P. Seeling, Eds., *Computing in Communication Networks - From Theory to Practice*, 1st ed., ser. 1. Elsevier, 2020, vol. 1.

[2] Hexa-X. (2021, Apr.) D1.2 – Expanded 6G vision, use cases and societal values — including aspects of sustainability, security and spectrum. [Online]. Available: https://hexa-x.eu/wp-content/uploads/2021/05/Hexa-X_D1.2.pdf

[3] ——. (2021, Dec.) D5.1 – Initial 6G Architectural Components and Enablers. [Online]. Available: https://hexa-x.eu/wp-content/uploads/2022/01/Hexa-X_D5.1_full_version_v1.0.pdf

[4] F. Bannour, S. Souihi, and A. Mellouk, "Distributed sdn control: Survey, taxonomy, and challenges," *IEEE Communications Surveys Tutorials*, vol. 20, no. 1, pp. 333–354, 2018.

[5] Arzo, Sisay Tadesse, Scotece, Domenico, Bassoli, Riccardo, Barattini, Daniel, Granelli, Fabrizio, Foschini, Luca, and Fitzek, Frank H. P., "Msn: A playground framework for design and evaluation of microservices-based sdn controller," *Journal of Network and Systems Management*, vol. 30, no. 1, pp. 1573–7705, 2021.

[6] ONOS, "micro ONOS," https://docs.onosproject.org/, 2020, [Online; accessed 01-May-2021].

[7] "3GPP TR 21.915 version 15.0.0 Release 15," https://www.etsi.org/deliver/etsi_tr/121900_121999/121915/15.00.00_60/tr_121915v150000p.pdf, [Online; accessed 01-May-2021].

[8] ETSI. (2021, Dec.) 2013 ETSI GS NFV 002 V1.1.1: Network Functions Virtualisation (NFV) - architectural framework. [Online]. Available: https://www.etsi.org/deliver/etsi_gs/nfv/001_099/002/01.01.01_60/gs_nfv002v010101p.pdf

[9] T. B. Meriem, R. Chaparadza, B. Radier, S. Soulhi, J.-A. LozanoLópez, and A. Prakash, "GANA - Generic Autonomic Networking Architecture," 2016.

[10] S. T. Arzo, R. Bassoli, F. Granelli, and F. H. Fitzek, "Multi-agent based autonomic network management architecture," *IEEE Transactions on Network and Service Management*, pp. 1–1, 2021.

[11] "Mobile Edge Computing A key technology towards 5G - White Paper," https://www.etsi.org/images/files/ETSIWhitePapers/etsi_wp11_mec_a_key_technology_towards_5g.pdf, [Online; accessed 01-May-2021].

[12] "System Architecture for the 5G System," https://www.etsi.org/deliver/etsi_ts/123500_123599/123501/15.03.00_60/ts_123501v150300p.pdf, [Online; accessed 01-May-2021].

[13] A. Tootoonchian and Y. Ganjali, "Hyperflow: A distributed control plane for openflow," in *Proceedings of the 2010 Internet Network Management Conference on Research on Enterprise Networking*, USA, 2010, p. 3.

[14] R. Bassoli, "Chapter 7 - network function virtualization," in *Computing in Communication Networks*, F. H. Fitzek, F. Granelli, and P. Seeling, Eds. Academic Press, 2020, pp. 119–132.

[15] C. N. Serugunda, S. T. Arzo, F. Granelli, R. Bassoli, M. Devetsikiotis, and F. H. Fitzek, "Autonomous network traffic classifier agent for autonomic network management system," in *2021 IEEE Global Communications Conference (GLOBECOM)*, 2021, pp. 1–6.

Sisay T. Arzo is a postdoc at the Univ of New Mexico, USA. He received his PhD and MSc. in Telecom Eng from Uni of Trento. He received his BSc. from Hawassa Univ., Ethiopia. He has five years of industrial experience in Telecom industry. His research interest included network softwarization, IoT and Network Automation.

Domenico Scotece is a postdoc researcher at the Univ of Bologna. He received his Ph.D. and M.Sc. degrees in Computer Science Eng. from Univ of Bologna in 2020 and 2014, respectivly. His research interests include pervasive computing, middleware for fog and edge computing, IoT, and management of cloud computing systems.

Riccardo Bassoli is a senior researcher with the Deutsche Telekom Chair of Communication Networks at the Faculty of Electrical and Computer Engineering, Technische Universität Dresden (Germany). He received his Ph.D. degree from 5G Innovation Centre (5GIC) at University of Surrey (UK), in 2016.

Fabrizio Granelli Fabrizio Granelli is Full Professor at the Dept. of Information Eng. and Computer Science (DISI) of the Univ. of Trento,Italy. He was IEEE ComSoc Distinguished Lecturer for 2012-15 and 2021-22, Director for Online Content in 2016-17 and Director for Educational Services in 2018-19. He is author of more than 250 papers on networking.

Luca Foschini is Ass. Prof. of computer eng. His interests span from integrated management of distributed systems and services to edge computing, from management of cloud computing systems to Industry 4.0. He is a voting member of the IEEE ComSoc EMEA Board.

Frank H.P. Fitzek received the Dipl.-Ing. degree in electrical eng. from the Rheinisch-Westfälische Technische Hochschule, Aachen, Germany, in 1997 and the Dr.-Ing. degree in Electrical Engineering from the Technical University Berlin, Germany. He is a Professor and head of the Deutsche Telekom Chair of Communication Networks at Technische Universität Dresden.