POLITECNICO DI TORINO
Repository ISTITUZIONALE

A unifying operating platform for 5G end-to-end and multi-layer orchestration

A Unifying Operating Platform for 5G End-to-End and Multi-Layer Orchestration

Antonio Manzalini[*], Diego R. Lopez[‡],Håkon Lonsethagen[¶], Lucian Suciu[**],Roberto Bifulco[‡‡],Marie-Paule Odini[xi], Giuseppe Celozzi[xiii],Barbara Martini[xv]
[*] Telecom Italia Mobile, Italy
[‡] Telefónica I+D, Spain
[¶] Telenor, Norway
[**] Orange, France
[‡‡] NECLabs, Germany
[xi] HPE, France
[xiii] Ericsson, Italy
[xv] CNIT, Italy

Fulvio Risso[†],Jokin Garay[§],Vassilis Foteinos[‖], Panagiotis Demestichas[††],Giuliana Carullo[x], Marco Tambasco[x],Gino Carrozzo[xiv]
[†] Politecnico di Torino, Italy
[§] Keynetic Technologies, Spain
[‖] WINGS ICT Solutions, Greece
[††] University of Piraeus, Greece
[x] CoRiTeL, Italy
[xiv] Nextworks, Italy

***Abstract*—Heterogeneity of current software solutions for 5G is heading for complex and costly situations, with high fragmentation, which in turn creates uncertainty and the risk of delaying 5G innovations. This context motivated the definition of a novel Operating Platform for 5G (5G-OP), a unifying reference functional framework supporting end-to-end and multi-layer orchestration. 5G-OP aims at integrated management, control and orchestration of computing, storage, memory, networking core and edge resources up to the end-user devices and terminals (e.g., robots and smart vehicles). 5G-OP is an overarching architecture, with agnostic interfaces and well-defined abstractions, offering the seamless integration of current and future infrastructure control and orchestration solutions (e.g., OpenDaylight, ONOS, OpenStack, Apache Mesos, OpenSource MANO, Docker, LXC, etc.) The paper provides also the description of a prototype that can be seen as a simplified version of a 5G-OP, whose feasibility has been demonstrated in Focus Group IMT2020 of ITU-T.**

I. Introduction

A number of drivers are steering the evolution of ICT and telecommunications infrastructures: among them the pervasive diffusion of fixed and mobile ultra-broadband, performance advances in chipsets, the tumbling costs of hardware, the large availability of Open Source Software, advances of Artificial Intelligence and Machine Learning, all coupled with new advanced terminals capable of unprecedented computational power.

The trajectories of these drivers are aligning with the trend, usually termed as *Softwarization*, through which ICT and telecommunications infrastructures are radically leveraging on virtualization technologies to implement the so-called Digital Business Transformation. Traditional ICT and telecommunications application scenarios are heavily impacted by *Softwarization of Networks (SwNets)*; but with the advent of 5G this SwNets approach is convincingly evolving to the vertical industries using the communication infrastructure (e.g., Industry 4.0, Precision Agriculture, Smart Cities, Robots, etc.), thus resulting in a key architecture principle needed to implement the foundations of the future networks.

In this respect, Cloud [1], Edge and Fog [2] Computing, Software Defined Networking (SDN) [3] and Network Function Virtualization (NFV) [4] are the most investigated enabling technologies and can be seen as different dimensions of an overall trend.

Despite the numerous research and development efforts in the area of SDN and NFV have been going on for many years, with a number of products now in the market and a significant steering role of large-scale open source development communities (e.g. those behind the developments of OpenDaylight, ONOS, OpenStack, Apache Mesos, OpenSource MANO, Docker, LXC, etc.), it is still difficult to find consolidated control and orchestration solutions that can be easily taken up by Telcos and service providers to implement end-to-end the various 5G scenarios for their vertical customers. For instance, SDN controllers lack common application interfaces (northbound Interfaces), NFV orchestrators rely on different infrastructure models, etc. The heterogeneity in the implemented solutions is heading for complex and costly situations, with high fragmentation, which in turn creates uncertainty and the risk of delaying 5G innovations.

This paper argues that 5G should rely on an Operating Platform (5G-OP) capable of handling the 5G infrastructure as a flexible and highly adaptable virtual environment of logical resources, executing any network functions and services as "applications". This paper introduces the concept of such a 5G-OP, describing its main characteristics and design principles, as highlighted by some of the most significant use cases for 5G.

Accordingly, the outline of the paper is the following. Section II presents the master guidelines we envision for the 5G-OP; Section III describes prototype software architecture which can be seen as a simplified version of a 5G-OP, and that has been demonstrated at the Focus Group IMT2020 of ITU-T meeting (Geneva, Dec. 5th - 9th, 2016); Section IV provides some closing remarks.

II. 5G - Operating Platform

5G-OP is defined as a reference functional framework aiming at integrated management, control and orchestration of computing, storage, memory, networking resources as well as of resources at the network edge (e.g. sensors/actuators in the IoT ecosystem) and resources at end-user devices and terminals (e.g., robots and smart vehicles). 5G-OP is an overarching architecture, with agnostic interfaces and well-defined abstractions, offering the seamless integration of current and future infrastructure control and orchestration solutions (e.g., ONOS [5], OpenDaylight [6], OpenStack [7] and even Robot Operating System [8], etc.).

The 5G-OP concept raises from the need of extending the "Software-Defined Infrastructure" concept beyond the SDN/NFV/Cloud infrastructure components while generalizing it with respect to mechanisms for resource and service virtualization, abstraction and slicing. Indeed, the 5G-OP offers novel orchestration mechanisms not only relying on existing infrastructure controllers for SDN (e.g., OpenDaylight, ONOS), Cloud (e.g., OpenStack) and NFV (e.g., Open Source MANO), but also considering and seamlessly integrating another set of controllers related to 5G radio (e.g., OpenAirInterface), edge devices (e.g. IoT frameworks) and even end-user devices and terminals (e.g., Robot Operating System).

5G-OP results in a cross-industry orchestration solution, in which the various technological domains unified under a common framework run as plug-ins and are offered to the infrastructure owners and various tenants to build their specific customization and value added services.

This approach guarantees ease of integration of infrastructure platforms, that along with the use of open-source software, will result the boost and quick exploitation of open innovations in a wide range of areas, spanning from resource management to third-party creation of vertical application services on 5G.

A. Unified service model and 5G abstractions

One of the main distinguishing characteristics, and most challenging aspects, of 5G-OP is the ability to seamlessly supports new capabilities and services, while internal entities can evolve independently. With respect to new capabilities, the 5G-OP must be able to support new technological domains without impairing the existing ones, and to handle new objects (e.g., a new type of IoT sensor) that may be available in the infrastructure. With respect to new services, we can envision the necessity to support new application-specific orchestrators, or in general any software module that can perform some advanced computation (e.g., analytics) out of the data generated by the infrastructure or provide new services (e.g., QoS in an SDN domain). This intrinsic extensibility enables 5G-OP to evolve, while still supporting existing services. Furthermore, it enables the exploitation of the peculiar characteristics that are available at the infrastructure level and that may be lost with an approach based on the minimum common denominator, which is typical for abstraction layers that aim at exporting a unified model that is consistent across different platforms.

5G-OP will not have to go through heavy changes in order to support the new capabilities, resources and services. This is achieved by simply adding new (software) modules, which are seamlessly integrated into the 5G-OP to handle the additional features thanks to the 5G-OP model-driven abstraction, which facilitates the service composition of abstracted entities.

This concept represents one of the unique characteristics of 5G-OP, which provides unified abstractions and models that can be consistently used by all the orchestration services, running at any level for the continuous on-boarding of new capabilities, resources and services, without affecting any already active service instance, across various technologies in different administrative domains (i.e. with technology agnostic and federation mechanisms) and by allowing new services to use the new features (i.e. plug-and-play approach).

B. A Generalized Orchestration Space

The problem of orchestrating infrastructure-level services, i.e. the ones that need to be mapped on physical resources, is only a part of a bigger orchestration problem. Indeed, additional service orchestrators exist on top of an abstracted platform, which optimize the deployment of application-layer services, such as a Hadoop service running on Apache Mesos, which is hosted on an OpenStack-managed datacenter.

Through the definition of a "shared orchestration space"(shown in Figure 2), the 5G-OP brings together two problems that are usually considered separately: infrastructure-level and application-layer orchestration.

The ambition for generalized orchestration originates from the fact that the 5G-OP includes everything spanning from the end-user terminals to the core network and datacenter, including all the software layers running on all the above devices, thus also addressing application services.

To this direction, a generalized orchestration workflow/process should be devised that involves the composition of both application and infrastructural resources, capabilities and services while adapting the composite services to different and/or ever-changing contextual information [9].

With the aim to achieve a model-driven provisioning of services through the different levels of orchestration, one 5G-OP key feature is the availability of a common data model based on graphs that correlates and connects services, resources and capabilities together to represent relationships and workflows, as shown in the example in Figure 1.

In particular, a graph-based model, where services at a given layer are mapped onto services, resources and capabilities abstracted from the underlying layers, will enable a set of transformations/verification processes, which can be built on a solid mathematical groundwork based on graph theory. The service orchestrators will use a reference set of operations on graphs to be applied at different levels of abstraction, with a formal description of interfaces and expected graphs.

By these means, the relations between orchestration modules could be defined as a set of transformations and be formally verified. In order to improve performance and scalability of services, the transition across multiple layers could be

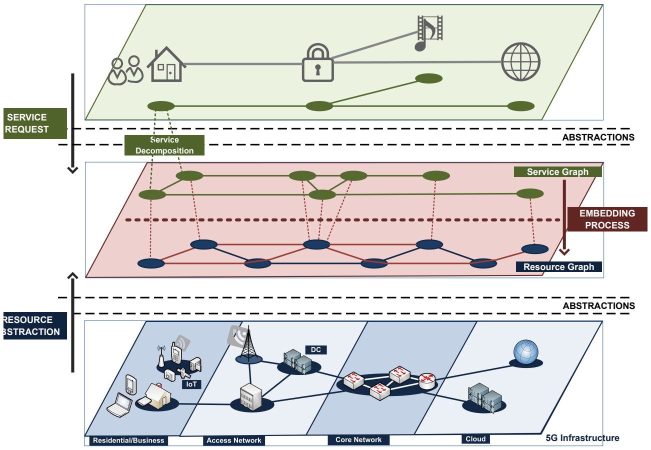

Fig. 1. Service and resource graphs.

optimized through an equivalent and formally verified graph-based model, enabling the definition of an allowed set of transformations. In this way, at runtime it will be possible to involve in the orchestration workflow just the service and infrastructure orchestrators strictly needed and provide a sort of "fast path" for the deployment and management of services. At the same time, formal verification of transformations can be leveraged as a way for ensuring the correctness of orchestration with respect to general or security-oriented policies (e.g. isolation properties).

This consistent approach between design and runtime phases will allow 5G-OP to reduce capacity churn, eliminate isolated under and unused capacity, reduce dependability and security issues, and respond to service requests in a sustainable, efficient and effective manner delivering the best user experience.

C. Main Architectural Principles

The 5G-OP allows a generalized, flexible and de-structured orchestration workflow in which orchestrators can decompose a service request into more elementary ones, discover which entities are available that can serve the new service requests (making use of the advertised resource and capabilities), and finally map them to the best entities given a possible set of constraints including geographical location, QoS and security requirements. Flexible service decomposition is allowed by the possibility of orchestrators to arbitrarily and directly interact with one another. The decomposition process originates a workflow of service invocations (modeled as a dependency graph) that is specific for the given request, since it depends on (*i*) the originating intent (a.k.a., service request), (*ii*) the state of the system, (*iii*) the actual constraints associated to the given service (e.g., configuration parameters for QoS, traffic steering, etc.). The monitoring, collection, filtering and elaboration of the state of the system is a relevant part in 5G-OP to provide a truly orchestration that is able to dynamically adapt provisioned services to cope with context changes (e.g., different user's preferences or locations, data throughput degradation caused by network congestion, etc.).

The generalized orchestration is assured by proper abstractions and interfaces offered by orchestrators while interacting each other to address service requests in a structured service producers-consumers relationship. In 5G-OP, each service orchestrator exposes a Provider API for the NBI (North-Bound Interface), and a Consumer API for the SBI (South-Bound Interface). Composition is achieved by attaching a Provider API to a Consumer API, thus providing the additional advantage of allowing horizontal composition, not requiring strict vertical hierarchies.

Indeed, different layers of abstraction for network programming and configuration are possible in 5G-OP, in order to support in a more generalized way various different technology domains, type of resources and possible services. More specifically, the Provider API can offer different logical views of the underlying resource and service capabilities (for network and non-network parts) to the service consumers, thus realizing the slicing concept. The 5G-OP Provider API heavily supports the concept of intents to ease the way a service consumer can request a service from the underlying layer, ignoring technological details on how the actual resources are configured and the service provisioned.

At the Consumer API, abstraction is mainly aimed to wrap details of different devices and resource in the underlying layer, controlled as objects with generalized capabilities across various technology domains (i.e. from legacy devices to OpenFlow-based switches, to 5G radio terminals, IoT sensors/actuators, etc.). In addition, unified protocols and communication paradigms (e.g., publish/subscribe for capability/resource advertisement, client/server for service invocation and data queries) will be used in the interactions between the different entities, hence offering to programmers an abstract communication model that will be automatically implemented by the system.

In this sense, 5G-OP advances the prior-art of some H2020 relevant projects such as SONATA and 5G Exchange (5GEx). In fact, a main difference with respect to the SONATA architecture is the concept of "generalized" orchestration space (orchestrators communicate/interact with certain communication primitives such us pub-sub) which is beyond the traditional layering approach; moreover this "generalized" orchestration space is "agnostic" with respect to other available orchestration and control solutions available today or tomorrow. Still the concept of "generalized" orchestration space, highly distributed up to the terminals, is rather different from the 5GEx software architecture which is mainly aiming at cross-domain orchestration of services over multiple administrations or over multi-domain single administrations.

III. Prototype Demonstration

This section reports the brief description of a prototype software architecture show in Fig. 3 that can be considered an initial and simplified version of a 5G-OP. The prototype architecture, based on the open-source FROG orchestra-

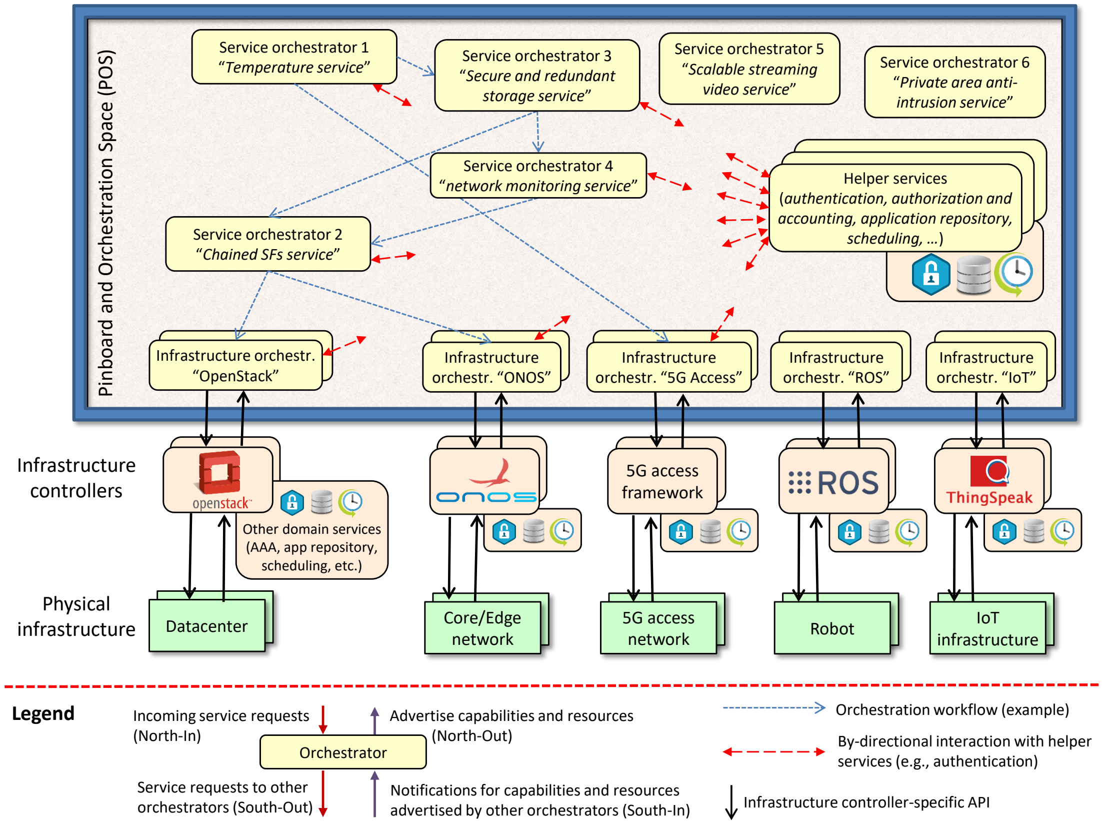

Fig. 2. 5G-OP Architecture.

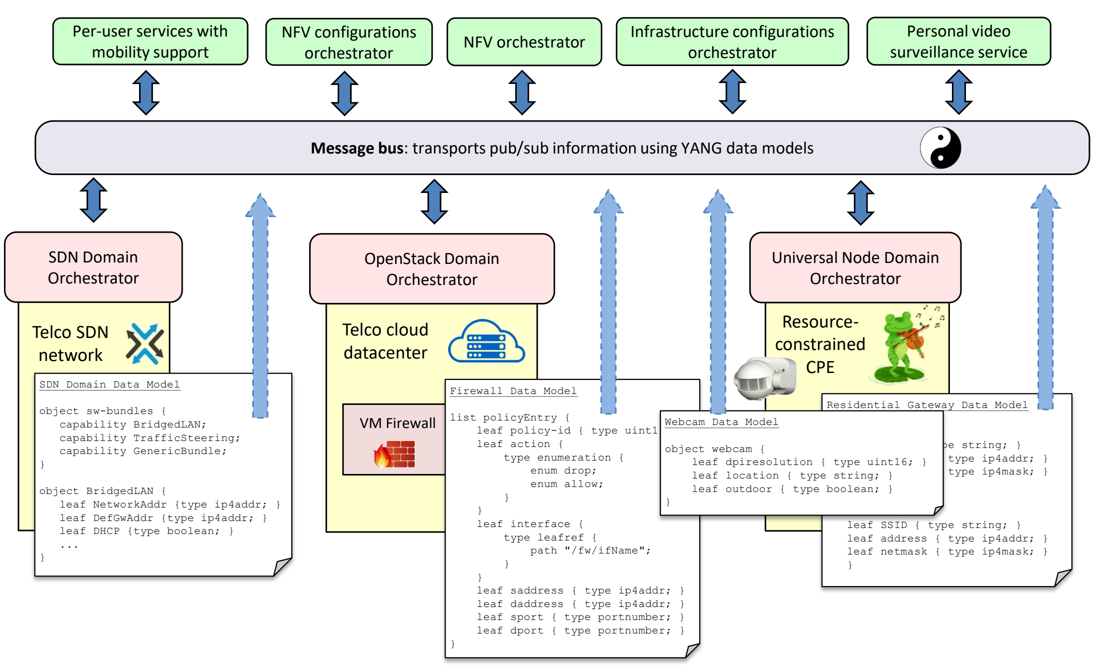

Fig. 3. Prototype demonstrated at the Focus Group IMT2020 of ITU-T meeting [10], [11].

tor [10], relies on a continuous advertisement of capabilities and resources from underlying infrastructure-layer domains, which allows the orchestration to adapt its service logic to exploit the most up-to-date capabilities. The advertisement process exploits a message bus that connects different types of entities, such domain controllers (e.g., tiny software layers that provide the interface between unmodified infrastructure controllers such as OpenStack and the message bus), network functions (e.g., a firewall), and individual resources (e.g., a sensor). Each of the above entities periodically advertise their capabilities/resources, while services (i.e., the ones on top of Fig. 3) receive the immediate notification whenever a capability/resource they are interested in (i.e., they subscribed for) has changed.

The YANG language has been selected to provide a unique data model for the above data exchange, hence providing a uniform common ground among all the entities. Direct interactions between entities are possible by means of a REST interface, which can be dynamically created based on the YANG model of the object itself and that supports the basic CRUD (Create, Read, Update, Delete) operations.

The prototype demo, which has been demonstrated at the Focus Group IMT2020 of ITU-T, showed the setup of a complex NFV service across multiple domains, such as user terminal (laptop) attached to an SDN network and asking to be connected to the Internet through a NAT; an OpenStack instance, connected by an SDN network, was available to execute possible network functions as virtual machines.

In the first part of the demo, the intermediate SDN network advertised only traffic steering capabilities, hence the FROG overarching orchestrator had to connect the user terminal to the Internet by steering the traffic to the data center where a NAT, available as a virtual machine, was launched; hence the intermediate SDN network was used only to connect all the different components together. However, when the SDN network advertised also the capability to host a given set of applications (e.g., a NAT), the overarching orchestrator adapted its service logic and it instantiated the entire service in the SDN domain (e.g., as ONOS applications), leaving the datacenter to host possible other services that may be requested in the future and that are not supported by the SDN domain.

Albeit simple, this prototype demonstrated the possibility and the advantages, for an overarching orchestrator, to change its behavior based on the prompt advertisement of capabilities/resources coming from the underlying infrastrcture. This can enable more aggressive optimization strategies, as well as a more effective (and timely) use of the available resources. For further details, see [10] and [11].

IV. Conclusions and Next Steps

Despite the efforts of standardization bodies and large-scale open source development communities, it is still difficult to find consolidated control and orchestration solutions that can be easily taken up by telcos and service providers.

The heterogeneity in the implemented solutions is heading for complex and costly situations, with high fragmentation, which in turn creates uncertainty and the risk of delaying 5G innovations. Moreover, it is not predictable today which of said platform(s) will be widely accepted and deployed, and how they will evolve. This context motivated the definition of 5G-OP as unifying Operating Platform for 5G end-to-end and multi-layer orchestration.

5G-OP is not another control-orchestration platform, at the level of the ones that are around today. On the contrary, 5G-OP is positioning above them with proper interfaces, universal set of abstractions and "adaptation" functions. The agnostic and overarching characteristics of the 5G-OP will allow decoupling from the underneath control-orchestration platforms, which will become pluggable in 5G-OP. Therefore, 5G-OP is not adding another layer of complexity, but it is radically simplifying the integrations of current and future platforms, mastering the heterogeneity in space and time. This easiness of integration will allow network operators to exploit quickly the innovation in the network operations and the service provisioning areas/processes, as soon as this innovation is emerging.

The paper provided the overall description of the 5G-OP software architecture and prototype which can be seen as an extremely simplified version of a 5G-OP, whose feasibility has been demonstrated in the Focus Group IMT2020 of ITU-T.

Security and *scalability* will deserve a special attention in our future studies. With respect to the former, the message bus becomes the nervous system of the entire architecture, hence must be able to preserve its operations also in case of attacks and provide a strong isolation between the different actors, as all the messages are transported across the same infrastructure. With respect to the latter, the definition of an architecture that scales at the geographical level, with hundred of millions of connected entities and still allow arbitrary communication between any entity, is definitely a challenge.

References

[1] S.-C. M. Inc, "Twenty experts define cloud computing," http://cloudcomputing.sys-con.com/read/612375_p.htm, 2008.
[2] A. Manzalini, R. Minerva, F. Callegati, W. Cerroni, and A. Campi, "Clouds of virtual machines in edge networks," *IEEE Communications Magazine*, vol. 51, no. 7, pp. 63–70, 2013.
[3] D. Kreutz *et al.*, "Software-defined networking: A comprehensive survey," *Proceedings of the IEEE*, vol. 103, no. 1, pp. 14–76, 2015.
[4] R. Mijumbi, J. Serrat, J.-L. Gorricho, N. Bouten, F. De Turck, and R. Boutaba, "Network function virtualization: State-of-the-art and research challenges," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 1, pp. 236–262, 2016.
[5] "The onos open source project," http://onosproject.org/.
[6] "The opendaylight open source project," http://www.opendaylight.org/.
[7] "The openstack open source project," http://www.openstack.org/.
[8] "he robot operating system (ros)," http://wiki.ros.org/ROS/Introduction.
[9] F. Paganelli, M. Ulema, and B. Martini, "Context-aware service composition and delivery in NGSONs over SDN," *IEEE Communications Magazine*, vol. 52, no. 8, pp. 97–105, 2014.
[10] "The FROG open source project," https://github.com/netgroup-polito/frog4.
[11] "ITU Workshop website," https://www.itu.int/en/ITU-T/Workshops-and-Seminars/201612/.