

Flexible Fog Computing and Telecom Architecture for 5G Networks

Luis Velasco* and Marc Ruiz

Universitat Politècnica de Catalunya (UPC), Barcelona, Spain

**e-mail: lvelasco@ac.upc.edu*

ABSTRACT

We review a novel, secure, highly distributed and ultra-dense fog computing infrastructure, which can be allocated at the extreme edge of a wired/wireless network for a Telecom Operator to provide multiple unified, cost-effective and new 5G services, such as Network Function Virtualization (NFV), Mobile Edge Computing (MEC), and services for third parties (e.g., smart cities, vertical industries or Internet of Things (IoT)). The distributed and programmable fog technologies are expected to strengthen the position of the Mobile Network and cloud markets; key benefits are the dynamic deployment of new distributed low-latency services. The architecture consists of three main building blocks: a) a scalable node, that is seamlessly integrated in the Telecom infrastructure; b) a controller, focused on service assurance, that is integrated in the management and orchestration architecture of the Telecom operator; and c) services running on top of the Telecom infrastructure.

Keywords: fog computing, data analytics, 5G services.

1. INTRODUCTION

Telecom operators are deploying cloud computing and storage infrastructures [1], integrated with their Software Defined Networking (SDN) -controlled heterogeneous access and transport networks, in order to provide Network Function Virtualization (NFV), as well as Cloud Radio Access Network (C-RAN) and ultimately, Mobile Edge Computing (MEC) for future 5G services. However, deploying a private cloud infrastructure integrated with the telecom network presents a differential set of challenges due to the industry's inherent requirements for high-availability (5-nines), ultra-low latency, and complex networking (Ethernet, optical, wireless, etc.).

Aiming at fulfilling this vision of distributed services being offered on top of intelligent resources spread through the whole network, from the core to the edge, the authors in [2] proposed TelcoFog as an innovative architecture based on SDN, NFV and MEC, extending the Fog computing concept to the network operator. The basics and objectives of Fog computing (e.g., geographical distribution, low-latency applications, mobility, etc.) have been previously explained in the literature [3]. Fog nodes, located at the network edge, are required to support a number of heterogeneous services by providing computational, storage and networking resources. These services have their own requirements in terms of bandwidth, latency, reliability, etc. Moreover, the adoption of Fog computing brings the possibility that (cloud) resources allocated for each service may be distributed among a number of functions (generic virtualized functions), which are instantiated through selected Fog nodes. OpenFog [4] proposes a distributed architecture that solves use cases that cannot be successfully implemented using "cloud only" solutions or independent edge devices. The distribution of functions and the subsequent forwarding of the (control and data) traffic through those functions is referred to service function chaining (SFC) [5]. Within the TelcoFog, the SFC concept is adopted in fog computing, in order to offer added-value services in all the targeted scenarios (i.e., IoT, MEC and NFV).

This paper reviews the basic components of the novel TelcoFog system architecture and specifically focuses on its big data analytics framework.

2. TELCOFOG OVERVIEW

To make the above concepts reality, it will be necessary to provide a control and orchestration system (the TelcoFog controller) being able to allocate cloud resources for the VNFs of a specific service along with defining its SFC. For the latter, the TelcoFog controller should also ensure the connectivity (network resources) among the different Fog nodes, accommodating the VNFs, to enable traffic steering according to the computed SFC. This provides a seamlessly and unified control for the complete visibility, computation and allocation of both cloud/fog and network resources through different network segments (access, aggregation and transport) assuming heterogeneous access and transport technologies (e.g., Wi-Fi, packet switching, optical transmission, etc.) with the goal of satisfying and ensuring (stringent) service requirements (e.g., zero-latency applications for mission-critical machine communications). To this end, a fundamental action to be done, currently not completely accomplished, is the definition of a common data model, for devices and TelcoFog services that will be used by TelcoFog controller.

Besides the Fog-like nature of TelcoFog and its orchestration capabilities, another critical aspect is the adoption of YANG as a data modeling language, which has been steadily growing in the IT and networking communities for the last few years, and consequently its use has been increasing in multiple standard defining

organizations – there is still a lot to define regarding computational resources and devices to cover the IoT landscape, from sensors to IoT gateways and related applications (such as IoT services and the enabled application ecosystem that better exploit the gathered knowledge) [6]. Moreover, it shall include smart and expert systems able to apply newer approaches involving collecting and processing large volume of data.

Another important aspect to standardize are Application Programmable Interfaces (API) between functional elements in the encompassing frameworks and related service workflows, including, as important ones, resource, device, service exposition and (auto)discovery.

The adoption of a common, flexible and powerful data and information modeling language to express all sensors, actuators, gateway facilities and services is a first, important step towards the standardization of IoT frameworks across multiple vendors beyond the existing ones. Moreover, a standardized IoT framework shall include the optimization of the underlying data transport network. Since this objective is known to be ambitious and mid to long term, an important goal is also ensuring at least a certain level of interoperability in complex deployments. The widest adoption of a common set of models, methodologies and approaches is an enabler for a further automation in the integrated control and management of IoT platforms that include, in addition to the main components in an IoT deployment, the integration of the heterogeneous transport network as a core component.

The above opens a whole new set of significant challenges:

1. *Fog computing for Network Operators.* Cloud interconnection has been studied, but joint integration of fog and cloud paradigms and their network interconnection remains a research challenge
2. Fog and cloud computing interconnection through operator SDN-enabled networks.
3. *Decentralized NFV services.* NFV enables the migration into a telecom cloud and the transition of the operator network to SDN control, allowing at the same time the programmability of the network resources for fully utilizing the capacity of the deployed physical resources.
4. *Mobile edge computing and D2D communication.* Using MEC, network operators can deploy fog nodes in the RAN and open the radio network edge to third-party partners, allowing them to rapidly deploy innovative applications and services towards mobile subscribers, enterprises and other vertical segments
5. *Integration with a monitoring and data analytics framework.* The challenge is to provide the necessary extensions to big data frameworks in order to provide the necessary TelcoFog node computational and storage resources for applications.

Figure 1 provides an overall description of the basic components. TelcoFog nodes are placed at the edge of the network. TelcoFog nodes might create an overlay network (referred as TelcoFog edge network) on top of the operator's access network in order to interwork allowing different networking services (such as multicast or anycast services based on the emerging segment routing technology). In both scenarios, the TelcoFog node plays a significant role in providing the necessary distributed and secure computing, storage and networking infrastructure in order to provide the targeted services. The TelcoFog controller can both, act in a hierarchical (parent/child) or peer architecture for the optimal allocation of resources across several networks and TelcoFog nodes. It also includes a mechanism for slicing such TelcoFog resources, as well as providing a standardized procedure to describe and deploy TelcoFog services.

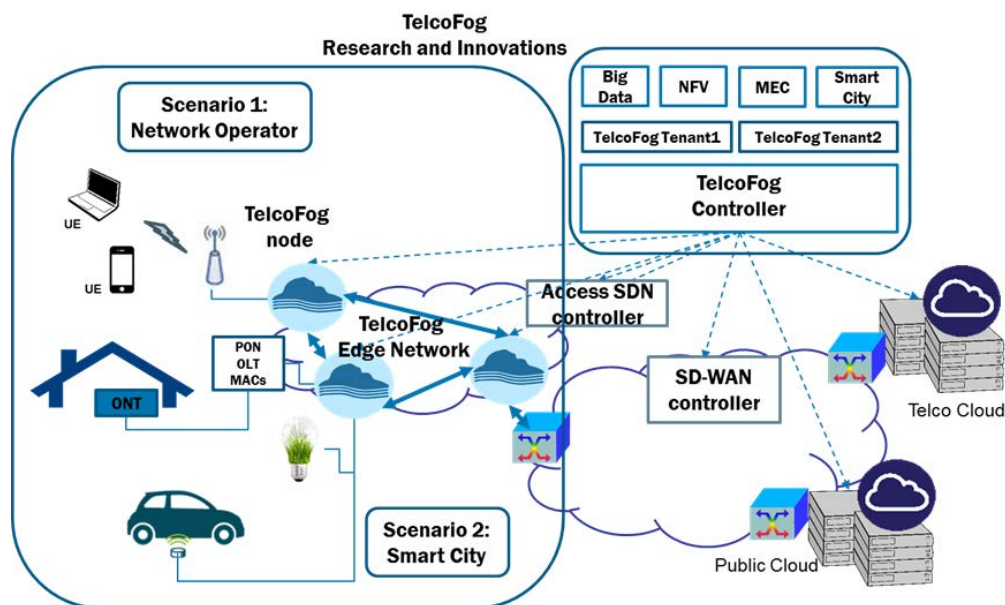


Figure 1. TelcoFog architecture (from [2]).

Finally, to complement these mechanisms, a big data analytics framework is also provided. One of the main objectives of data analytics is knowledge discovery from data (KDD) that includes data collection, pre-processing and storage, data transformation, and data mining and presentation.

Resource Monitoring is key in providing Service Assurance, since it inspects the provided TelcoFog services in order to anticipate future degradation of the Quality of Experience. In addition, an important component within the TelcoFog node are the Data Analytics Helpers, a predefined set of functionalities for data analysis and evaluation can be provided as services running on top of allocated virtual resources.

TelcoFog supports the KDD process and offers *knowledge-as-a-service* to TelcoFog services. The real-time capability will provide mission critical business intelligence in the support of enterprise decision making. Such framework will allow applications to monitor and manage computation and storage resources, while protecting the privacy and integrity of data.

3. TELCOFOG BIG DATA ANALYTICS FRAMEWORK

The architecture proposed in TelcoFog fits with the steps defined for the KDD process. The real-time capability will provide mission critical business intelligence in the support of enterprise decision making. Such framework will allow applications to monitor and manage computation and storage resources, while protecting the privacy and integrity of data.

Each component in the fog computing architecture generates detailed event records for every significant event. These event records are sent to one or multiple instances of the analytics node that collate and store the information in a horizontally scalable database based on e.g., Apache Hadoop, which uses a format optimized for data stream mining and queries. The analytics nodes include mechanisms to automatically trigger the collection of more detailed records when certain event occurs; the goal is to be able to get to the root cause of any issue without having to reproduce it.

Built on Apache Hadoop distributed computing architecture, Apache Spark is a fast and general open source engine that is quickly climbing towards the top of big data processing frameworks. Spark offers the possibility to build parallel applications taking advantage of high-level operators to facilitate such development task. Spark can run in several modes, including the Yet Another Resource Negotiator (YARN) mode, which supports security, allowing running against Kerberized clusters and using secure authentication between processes. In the YARN model several key elements are defined to achieve a fully distributed computing. An application might consist of a number of jobs, a driver and an application master. Each application can require several executors distributed in many nodes. A highly distributed and ultra-dense Fog infrastructure enables multiple deployments of YARN-based computing clusters. The framework exposes northbound APIs that applications use to effectively leverage the Fog platform. These APIs are broadly classified into data and control APIs: data APIs allow an application to leverage the Fog distributed data store, whereas control APIs allow an application to specify how the application should be deployed on the Fog platform. In this way, the observe, analyze, and act cycle can be implemented.

A big data analytics provisioning system will be available for services requiring specific requirements, such as response time, area coverage, cost, etc. The system will generate provisioning configuration for the big data analytics framework. The real configuration will be managed by each application, so resources can be scaled in/out to adapt the application footprint to its current need in every geographical area.

4. SMART CITY USE CASE

The proposed architecture enables a hierarchical KDD and decision-making platform especially important for those latency-sensitive and geo-distributed applications, which are common in the smart city environment (see Fig. 2). Bottom-up, several levels are defined each carrying out its own KDD process to support decision making at that level. Note that applications within the same level can look at different time scales. For example, at the Sub-city level, the Green Wave Control works in the time scale of few tens of seconds, whereas the Pollution Measurement typically generates results averaged on time scale of at least one hour.

The data produced at each level is processed, transformed, cross-correlated, and pushed up to the upper level, whilst decision making generates boundary conditions for the lower level. Therefore, lower level applications focus on their own observe, analyze, and act cycle looking at a local objective, while upper level applications look at global system performance objectives. Multiple levels could be mapped on one or more Fog nodes; physical architecture will depend on dimension of involved urban area. Furthermore, fog/cloud architecture is strictly connected to the computing distribution need among different nodes and to provide the needed commands to the different actuators with a very limited latency.

In the example, a sub-system is defined as a set of smart traffic sensors and actuators (traffic lights, maximum speed signals, direction signals, etc.). For instance, each traffic light can be equipped with sensors to measure the distance and speed of approaching vehicles and detect also the presence of pedestrians crossing the street in red. By processing that data in real-time, a traffic light can change its behavior to, e.g. slow down approaching vehicles to avoid a probable collision with pedestrians, swap lights in an intersection to leave the way free for an

emergency vehicle, activate specific alert signals in case of detection of floods. These low-latency decisions can require the coordination of several signals on the sub-system and between adjacent sub-systems.

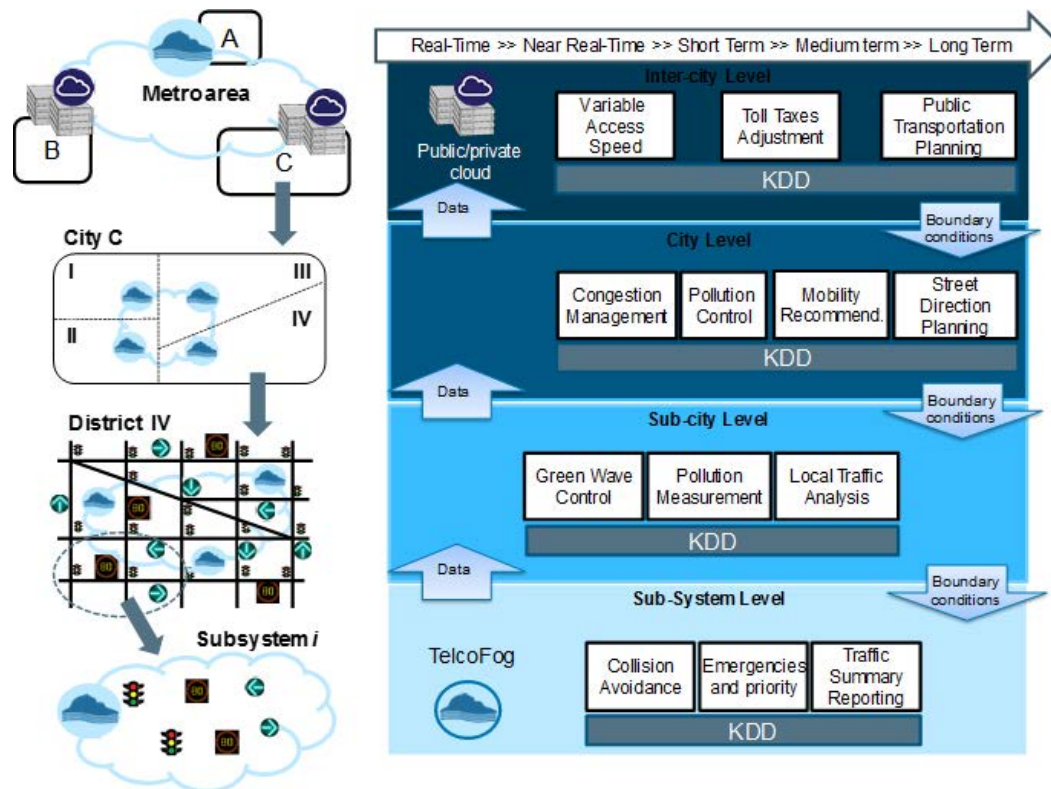


Figure 2. Hierarchical KDD in smart cities scenarios.

Above the sub-system level, the sub-city level groups several sub-systems that provide meaningful data to perform decisions oriented to, e.g. maximize speed within large city areas, while ensuring low air pollution. Note that cross-correlation is essential to obtain the best and most robust estimates from numerous data sources. The analysis of large data sets from different levels collected during long periods will allow city planners to improve city mobility, e.g., by modifying the direction of some main streets. On the top, an inter-city layer will allow optimizing those aspects related to metropolitan traffic including toll taxes policies to foster the use of some under-utilized highways, planning inter-urban public bus lines, etc.

5. CONCLUSIONS

In this paper we have reviewed a novel architecture for providing unified cloud and fog resources for deploying NFV, MEC and IoT service on top of a Telecom Operator's network. Specifically, the big data analytics architecture of TelcoFog has been described and a use case for smart city has been used to present the proposed hierarchical KDD process.

ACKNOWLEDGEMENTS

This work was partially supported by the EC through the METRO-HAUL project (G.A. n° 761727), from the AEI/FEDER TWINS project (TEC2017-90097-R), and from the Catalan Institution for Research and Advanced Studies (ICREA).

REFERENCES

- [1] L. Velasco *et al.*, "A service-oriented hybrid access network and cloud architecture," *IEEE Communications Magazine*, vol. 53, pp. 159-165, 2015.
- [2] R. Vilalta *et al.*, "TelcoFog: A unified flexible fog and cloud computing architecture for 5G networks," *IEEE Communications Magazine*, vol. 55, pp. 36-43, 2017.
- [3] F. Bonomi, R. Milito, J. Zhu, and S. Addepalli, "Fog computing and its role in the internet of things," in *Proc. MCC Workshop on Mobile Cloud Computing*, 2012.
- [4] OpenFog Consortium Architecture Working Group, OpenFog Architecture Overview, White Paper, 2016.
- [5] J. Halpern and C. Pignataro, "Service Function Chaining (SFC) Architecture", IETF RFC 7665, 2015.
- [6] Benoit Claise, "YANG as the data modeling language in the IoT space," in *Proc. IoT Semantic Interoperability Workshop*, 2016.