

# Supporting QoE/QoS-aware end-to-end network slicing in future 5G-enabled optical networks

Rafael Montero, Albert Pagès, Fernando Agraz, Salvatore Spadaro  
Optical Communications Group, Universitat Politècnica de Catalunya (UPC), Barcelona, Spain

## ABSTRACT

Network slicing with Quality of Experience/Quality of Service (QoE/QoS) guarantees is seen as one of the key enablers of future 5G networks. Nevertheless, it poses several challenges in both resource provisioning and management that need to be addressed for the efficient end-to-end service delivery. In particular, network slice deployments considering operation across several domains and network segments, require of inter-domain configurations, continuous monitoring, potential actuations, inter-slice isolation, among other, in order to be provisioned and maintained, looking forward to guaranteeing their assured Key Performance Indicators (KPIs). In such scenario, optical networks are of prime importance, enabling the inter-connectivity between multiple far away segments and Points of Presence (PoPs). In light of this, in this paper we present an architecture design enabling network slice provisioning for 5G service chaining in multi-segment/multi-domain optical network scenarios. The presented design is enriched with a policy-based monitoring and actuation framework able to maintain the desired QoS for the provisioned end-to-end (E2E) network slice. We experimentally validated the proposal against real slice deployments and traffic generation, providing a proof of concept for the presented architecture, with special emphasis in the demonstration of the actuation framework as a key element for quality guarantees.

**Keywords:** 5G, Network Slicing, Service Chaining, QoS Sensors and Actuators, Data Centers, Optical Networks

## 1. INTRODUCTION

As telecom and network infrastructures move towards the 5<sup>th</sup> Generation (5G) of mobile networks, new challenges appear on the arena in order to efficiently provision services over the shared network infrastructure. Not only future 5G networks will have to withstand substantially increased peak and average bit-rates per single user, increased user density and volume as well as reduced E2E latencies<sup>1</sup>, but also will need to support high levels of customization and flexibility to cope with the requirements of key stakeholders of the economic activity (e.g., eHealth, Automotive, Industry 4.0, Energy, etc.). Indeed, these 5G vertical customers define very specific requirements which have to be accommodated independently over the shared underlying infrastructure, providing also the means for customization and service isolation. In this regard, services in 5G are broadly categorized in enhanced mobile broadband (eMBB), ultra-reliable and low latency communications (URLLC) and massive machine-type communications (mMTC) service types, each of them imposing different KPI requirements. As a result, future network scenarios need to define solutions to efficiently cope with multi-tenancy in a graceful way.

Network slicing and, thus, network slices (NSes) have been identified as the enabling technology for efficient service delivery in future 5G networks<sup>2</sup>. In essence, NSes leverage onto the Network as a Service (NaaS) paradigm, in which a common network infrastructure can be leveraged by different verticals, partitioning the infrastructure in self-contained logical networks customized to the service to be delivered on top. In such a process, resource virtualization, either network or computing, becomes primordial. Thanks to virtualization, abstract representations of physical resources can be obtained and then composed in unified logical instances (i.e. the NSes). Such virtualization becomes the foundation for slice customization, as each virtual element can be specifically tailored to fit the needs of the slice. Furthermore, virtualization also enables the possibility to not only offer resources to compose the slice, but also Virtual Functions (VFs), that implement specific processes, such as traffic shaping, video encoding, firewalling, to name some, that enrich the provisioned NSes.

Software Defined Networking (SDN) and Network Function Virtualization (NFV) are the key enablers for network slicing in 5G<sup>3</sup>. On the one hand, SDN enables the abstraction of network resources, providing granular control to consumers. Abstracted resources can then be easily configured and monitored so changes in the network can be triggered

when required. At its turn, NFV paradigm allows for the creation and deployment of Virtual Network Functions (NFV), which encapsulate specific networking or application functions, to provide specific flavors for NSes customization. This is achieved thanks to the chaining of multiple VNF instances, interconnected through configured data paths in both virtual and physical network infrastructures.

In end-to-end scenarios, NSes will span different segments and technological domains (radio, wired access, core, transport), with NFV PoPs (e.g., data centers (DCs), Mobile Edge Computing (MEC)) distributed across to support the deployment of custom functions. Seeing this, optical networks and their virtualization are entertained to be part of the E2E NS delivery, interconnecting far apart network segments or constituting the internal fabric of PoPs to enable high speed, high performance and low latency communications across the endpoints of the slices. For this reason, architectures and techniques to provision NSes in NFV-based optical networks are a must and currently being heavily investigated<sup>4</sup>. On another hand, one of the key aspects in 5G network slicing relates to the delivery and maintenance of QoS/QoE guarantees of the provisioned NSes. Indeed, quality maintenance is a daunting challenge that involves several aspects, ranging from slice/service monitoring to detect fluctuations on the quality or changes in the requirements, to re-configuration/scaling operations (actuators) to modify the provisioned resources to support the desired quality levels. In light of this, proper solutions to address such challenges must also be found.

Taking all of this into consideration, in this paper we focus on achieving the E2E provisioning and QoS maintenance of NSes in a multi-segment/multi-domain optically-interconnected network scenario. To this end, we propose a design for an architecture capable of handling deployments, monitoring and actuators over the underlying segments of the network and the multiple NFV PoPs, with aims to automate the NSes deployment and quality assurance. Then, the proposed architecture is experimentally evaluated to assess the full deployment-monitoring-actuation cycle. The remainder of this paper continues as follows. Section 2 further elaborates on the design principles that enable slice composition, monitoring and actuation, paving the path for the actual architectural proposal. Then, Section 3 presents the overall E2E architecture including the components introduced in this paper to coordinate the NSes provisioning while guaranteeing their QoS agreements. Next, Section 4 presents the results in regards of the architecture evaluation. Finally, Section 5 summarizes the key achievements and conclusions of the paper.

## 2. END-TO-END NETWORK SLICE PROVISIONING AND MAINTENANCE

As introduced in the previous section, deploying and maintaining an end-to-end (E2E) NS requires coordination across network segments in both provisioning and runtime phases as well as a control and management stack to enable dynamic configuration and monitoring of the provisioned resources, thus being able to control the quality of the deployed NSes. Generally speaking, a NS is composed by a set of VNF instances interconnected with each other (what is known as the service chain), with the required virtual infrastructure (network and computing) to support the services that will run on top. Then, it is up to the management and control infrastructure of the multiple segments involved in the deployment to actually configure and monitor the resources to be able to deliver the E2E NS.

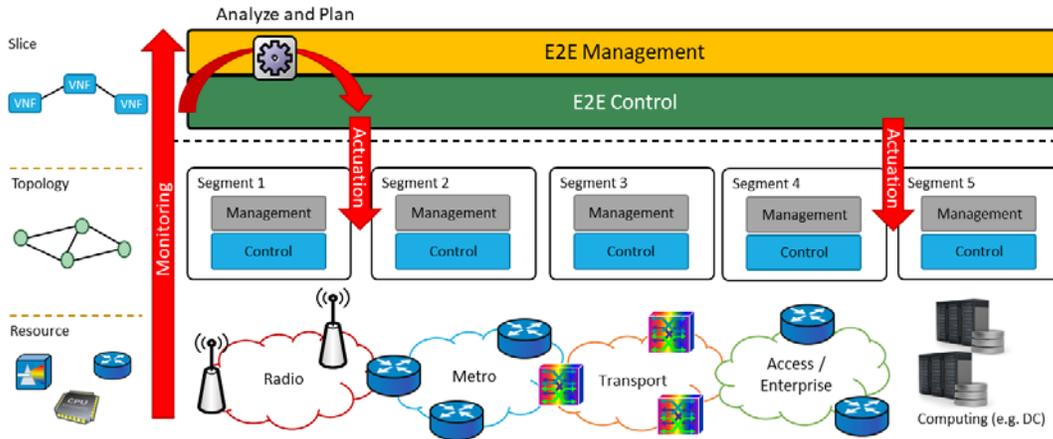


Figure 1. Schematic for MAPE-loop-based architecture for E2E slice control and management

To enable a holistic management, control and monitoring of the slices, an overlay coordination/management entity spanning the multiple involved segments is usually employed<sup>5</sup>. Such an entity is responsible to agglutinate all E2E control and management functions, distributing individual responsibilities to the involved segments once a proper resource layout has been determined for the NS to be deployed. Moreover, such an entity is also the responsible to collect monitoring information from the different segments to be able to construct higher level monitoring information of the E2E slice. Then, after analyzing the monitoring data, several re-configurations of the NS may be triggered through an actuation system overreaching the several segments in order to maintain desired QoS levels. Such loop is known as the Monitoring, Analysis, Planning and Execution (MAPE) loop and has become the foundation for autonomous and self-optimized networks (SONs). Figure 1 depicts a schematic of the loop, highlighting the main actors and interactions as well as the possible levels of monitoring.

Our architecture proposal follows such approach, leveraging on the capabilities of the multiple individual components to deliver a unified platform to close the full loop. The following sub-sections further elaborate on the main principles necessary for E2E slice provisioning with QoS guarantees, which are then adopted in the architecture presented in Section III.

### **2.1 Network Slices Composition**

One of the main challenges when delivering an E2E slice relates to the slice composition, that is, how the slice to support the wanted service is created and then partitioned in smaller slices to be deployed in the several segments. The Slice Composition concept (i.e., “slice-cum-slice”) introduced by the 5G-PPP<sup>1</sup> entails building a slice out of individual slices. In principle, different segments/domains could provide specific functionalities, then it would be possible to compose a slice by individually configuring parts of the E2E network infrastructure. As a result, a slice would be created from a joint set of functions, which are distributed across the several segments. Furthermore, a slice could also be extended by configuring additional functions and aggregating them to the slice.

More in particular, a NS Instance (NSI) then, would be instantiated by combining a set of NS Subnet Instances (NSSI)<sup>6</sup>. Generally speaking, each NSSI could belong to a different domain and be configured on a different network segment. A coordinating entity in such case would be required, so each NSSI is instantiated at the corresponding segment and inter-NSSI communication is also properly configured. Such level of slice composition would enable E2E slices to be built by provisioning and combining a number of NSSIs across all required network segments. Additionally, service-chaining related implementations could also make use of this technique as they may require functionalities to be present in different parts of the network. Finally, this technique could also help in accommodate NS deployments which functionalities or demanded resources are not available in a single segment/domain.

### **2.2 Monitoring and Sensors**

Considering the maintenance of the NS, it becomes crucial to keep up-to-date monitoring information of the underlying network and computing resources. Network sensors in turn, allow gathering specific monitoring QoS parameters from the network, so these can be furtherly aggregated to the NS level and processed. A sensor could be considered in many forms (e.g., Virtual Machine (VM), VNF, application) and be present at different levels of the architecture (i.e., infrastructure layer, control layer, orchestration layer). Moreover, some sensors may be instantiated and configured at the provisioning stage while others may be deployed on-demand during runtime. In this regard, such functional flexibility of sensors will help in the maintenance of a NS by putting emphasis only on monitoring specific network parameters related to guaranteeing its assured KPIs.

### **2.3 Policy-based Actuation**

After the collection and processing of monitoring data, either by using predictive or preventive techniques, and upon the possibility of the slice QoS of being compromised, actions (i.e. (re-)configurations) may be triggered to accommodate the slice to the current state of the network, thus still guaranteeing its required QoS. As in the case of the sensors, actions could be present in form of configurations at different levels of the architecture or may even consider deployments of a component (e.g., VM, VNF, application, sensor).

One popular approach to the actuation framework relates to the adoption of a policy-based actuation system, in which a policy framework creates policies tied to specific behaviors in regards of configurations needed at different levels of the whole infrastructure. Then, these policies are distributed across the multiple decision points, which are the responsible to determine when a policy has to be applied or not. A policy, thus, has to define under which conditions is activated and what are the resulting outcomes of applying the policy (i.e. the actions/configurations). In this regard, the Simplified Use

of Policy Abstractions (SUPA) working group initiative at the IETF defines an information model to represent different types of network policies<sup>7</sup>. A common SUPA model is composed by an event (e.g., packet drop ratio), a condition (e.g., greater than 10%) and an action (e.g. increase network bandwidth), which, in combination, dictate which are the triggers of the policies and what are the actions that need to be enforced when the conditions are met. Based on this, data models can be defined to serve as an input for network management functions, possibly resulting in network configuration changes under several conditions. These models would allow setting policy rules including commands to be sent between entities, as well as serving for managing and configuring devices and services.

### 2.4 NFV Coordinator

As a final principle, the work presented in this paper recognizes the need for a component capable of managing the slice composition while also being able of handling network sensors and actuators. Such a piece would then become the center of the whole MAPE loop for E2E NS delivery with QoS guarantees. As a result, an NFV Coordinator (NFV-C) comes to the picture, which would be the responsible for the following functionalities required for the provisioning and maintenance of NS:

- Coordination across single NFV entities to compose and provision slices.
- Inter-domain configuration to enable connectivity between segments allocating NSSIs.
- Inter-segment configuration to allow routing/chaining between VNFs across different NSSIs.
- Configuration and deployment of network sensors and monitoring data collection.
- Processing and analysis of gathered monitoring data to feed the actuation framework.
- Triggering of actuations to guarantee overall QoS of the slice.

## 3. ARCHITECTURE FOR E2E SLICE PROVISIONING

Following the concepts introduced in previous section, Figure 2 depicts the overall architecture presented in this work for multi-domain slice provisioning and monitoring in 5G-ready optical networks.

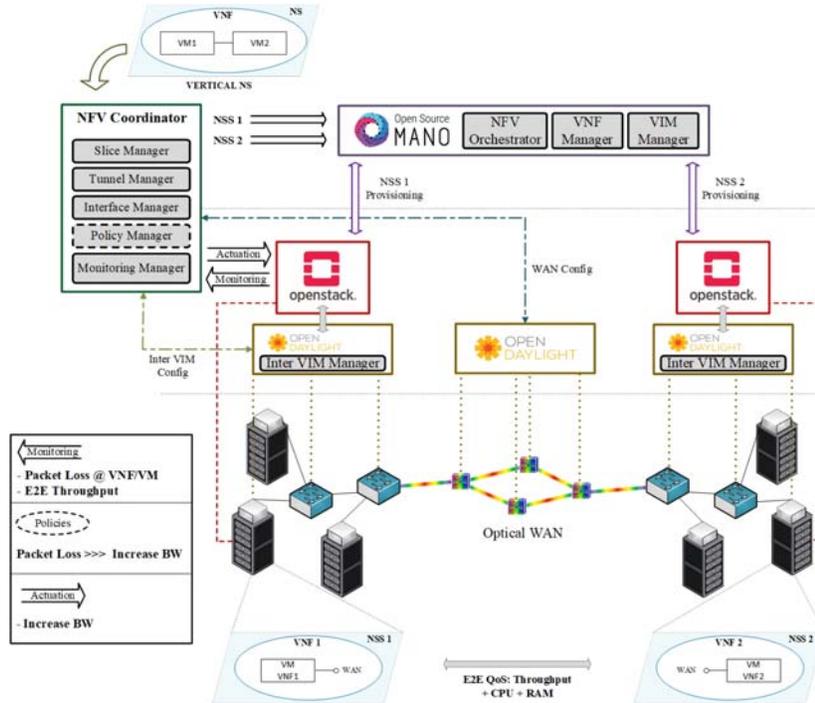


Figure 2. Architecture for E2E NS deployment in optically interconnected DCs.

As seen in the figure, we consider a multi-segment scenario with optically interconnected DCs through a Wide Area Network (WAN), with SDN controllers to configure network elements at the infrastructure layer. Elements belonging to DCs and optical network segments expose their functionalities towards an SDN controller (i.e., OpenDaylight (ODL)) via a southbound protocol. In turn, the controller offers network configurability to consumers in upper layers (e.g., orchestration, application). An orchestration entity (i.e., OpenStack) at the DC segments is in charge of managing computational resources and requesting the SDN controller for network connectivity. At the WAN segment, the request for network configuration is relegated to the NFV-C component. The NFV-C takes a major role in the whole architecture, as it not only helps in the provisioning of slices by partitioning them using the slice composition technique and requesting their allocation, but also allows for their ongoing monitoring and policy-based maintenance. A more in-depth characterization of this component follows at the analysis of the slice lifecycle stages. Finally, an NFV Management and Orchestration (MANO)<sup>8</sup> entity (i.e., Open Source MANO) is also required to manage the provisioning of per-segment NSSIs, that is, the deployment and configuration of VNF instances belonging to each NSSI.

### 3.1 Slice Provisioning Phase

The NFV-C acts as the start point for slice provisioning. It allows the consumer (e.g., a 5G vertical) to request for a network service based on a Network Service Descriptor (NSD). The Slice Manager component at the NFV-C handles the request, determines the NSI to be created and de-composes it into several NSSIs according to service and resource availability at the different network segments. After defining the NSSIs, the next step is to request to the MANO framework for their provisioning by using individual descriptors for each of them and instructing the selected segment for their allocation. The MANO entity then is able to request to the corresponding DC segments via their orchestrators to provision computational (i.e., VMs) and network resources for each NSSI. While these are being provisioned, the NFV-C will request the SDN controller at the WAN segment for network connectivity between segments, so the Tunnel Manager can trigger the creation of a tunnel between involved DCs for E2E communication across VNFs. Thanks to the optical capabilities, multiple tunnels are aggregated at border nodes in high bandwidth optical communications, enabling for efficient resource sharing across individual communications of a single NSI spanning multiple DC segments. Once VMs/VNFs are operative at both network segments, the NFV-C Interface Manager will be able to retrieve interface information at both ends and configure the inter-segment routing. The Inter-Virtual Infrastructure Manager (VIM) Manager application at the multiple SDN controllers will enable coordination with the NFV-C for these final configurations. After all is set up, VMs belonging to each VNF should be able to reach each other via the data network, thus completing the provision stage of the NS.

### 3.2 Runtime Slice Monitoring and Actuation Phase

Once the NS has completed the provisioning stage, the Monitoring Manager at the NFV-C starts an ongoing monitoring task particularized to the needs of the slice. In particular, it configures the data sources at the network segments, for instance, throughput and packet loss ratio for the intra-DC networks or bit error rate and Optical Signal to Noise ratio (OSNR) for the optical network segments, as well as monitoring sources for the computational resources, for example their CPU and RAM utilization. As said before, these data source serve as triggers for the whole actuation framework. In this regard, the Policy Manager component enables policy-based actions (i.e., actuations) that will allow maintaining the slice QoS agreements if compromised levels are reached. Policies are installed in the Policy Framework, either thanks to an automatic generation procedure (for example, based in machine learning) or manually on-boarded by a system administrator. These policies then dictate when and how actuations should be triggered according to the data sources for which they are bound to. When applying a policy that has been triggered, the NFV-C executes the required action, which can result in reconfigurations in both the computing and network resources provisioned for the E2E slice. As a result, a new infrastructure configuration is delivered thus guaranteeing the overall QoS of the slice.

## 4. EXPERIMENTAL VALIDATION

In this section, we experimentally validate the proposed architecture by providing a practical example of the whole monitoring and policy-based actuation loop. To this end, an emulated testbed consisting on 3 network segments (i.e., left-DC, WAN, right-DC), the same as depicted in Figure 2, is employed. In regards of control, each network segment is connected to its own ODL-based SDN controller, while each DC is managed by a single OpenStack entity. An overreaching OSM entity enables the deployment of NSSIs at particular DCs as result of the slice composition performed at the NFV-C element. For the particular tests, a single NS composed of two VNFs has been deployed, resulting in two NSSIs, each one with one VNF, deployed in different DCs. That is, as a result of the slice composition,

each DC hosts a single VM corresponding to the E2E NS, while network connectivity is delivered across the 3 involved network segments. To exercise the policy-based actuation, we focus on the case in which QoS of the slice is determined by the packet loss ratio experimented by the data flows across VNFs, which could be the case for video delivery services. To this end, a policy stating that as a result of a packet loss ratio (PLR) greater than a pre-defined threshold, the configured slice network bandwidth has to be increased to suitable levels, has been manually onboard to the system. PLR statistics are then collected at the virtual interfaces of the VNF instances. Given this setting, traffic of variable bit-rate is sent between the VNFs (in particular, the Iperf tool is employed) during a set amount of time and both detected throughputs (at sender and receiver) as well as PLR statistic are analyzed. Figure 3 depicts the obtained results as a function of the time in two situations: one in which an actuation framework is not present, thus no QoS guarantees, and another in which the whole actuation framework is exercised. For the two cases, the initial network capacity for the tunnel across the VNFs has been set to 25 Mb/s.

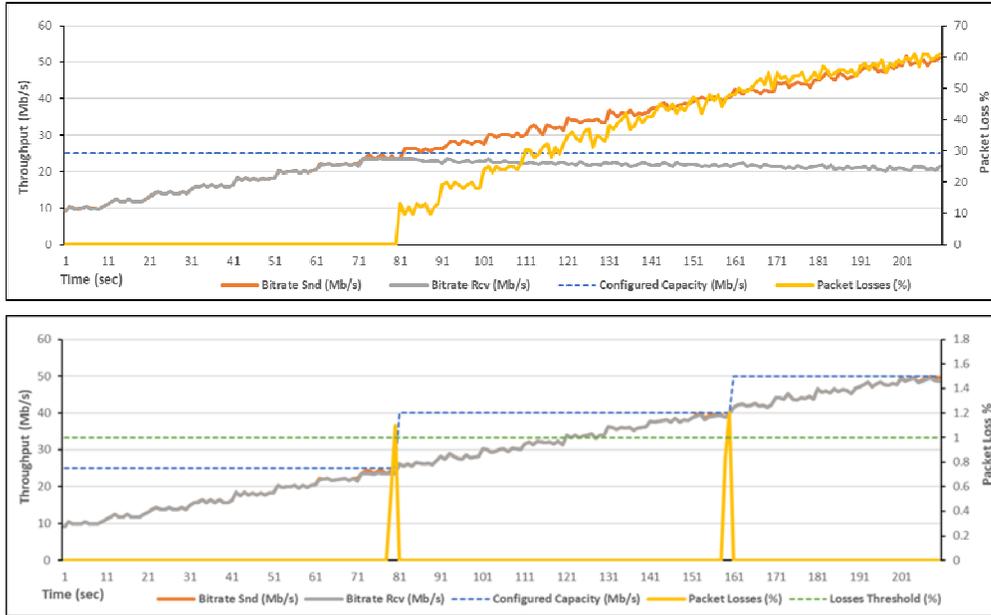


Figure 3. Throughput vs. Packet Loss tests without actuations (top) and Throughput vs. Packet Loss tests with policy-based actuations (bottom)

Figure 3 (top), depicts the case in which no policy nor actuation was set to overcome any detected packet loss. The throughput in this case, starts increasing from an initial value of around 10Mp/s until it reaches the 25Mp/s boundary (represented by the blue dashed line). As no preventive action is configured, the bitrate received at the server side reaches its limit and is maintained beneath the available bandwidth while the bitrate sent from the client side continues to increase and surpasses the limit, thus generating packet losses. As a result, when the bitrate sent by the client reaches 50Mp/s, the packet loss rate can get close to 60%, becoming an unsustainable situation in regards of QoS guarantees. This evidences the need of an actuation system to react to dynamic slice/service requirements and maintain desired quality levels. Then, we demonstrate how the Policy Manager module is able to handle policy-based actuations to overcome low network performance (Figure 3 (bottom)). In this case, as before, throughput starts increasing gradually from 10Mb/s. As it approaches the 25Mp/s capacity limit, the Monitoring Manager starts detecting packet losses. Once the threshold (1% is used for the experimental validation) is reached (represented by the green dashed line) the policy is activated and the Policy Manager demands for an increase of bandwidth at the provisioned network resources to overcome the packet losses. As shown, the new limit is set to 40Mb/s while the percentage of lost packets drops again to zero. In this case the throughput continues to grow, reaching again the new limit. As new packet losses are detected and the preventive threshold is exceeded again, the policy-based actuation is triggered a second time, increasing the provisioned bandwidth to 50 Mb/s. Packet Loss ratio drops one more time to zero, thus maintaining the performance of the network at the required standards. In this regard, it can be appreciated how the proposed policy-based actuation framework is capable of maintaining desired quality levels over the time under dynamic conditions.

## 5. CONCLUSIONS

With the rise of modern network scenarios for 5G deployment, techniques such as Network Slicing, Slice Composition and Service Chaining, among others, come along to leverage legacy networks towards higher demanding requirements. In this work we presented an architecture capable of handling the provisioning and maintenance of 5G services through the use of such techniques in optically interconnected network scenarios. In turn, a Network Slice can be de-composed for its deployment across different segments of the network, considering the required inter-VIM configurations by the introduced NFV-C component. Furthermore, services can be maintained through the ongoing monitoring of specific network parameters and by operating with a policy-based actuation behavior. Experimental results shown in the paper demonstrate how required performance levels of the service can be guaranteed in case of adverse QoS performance to dynamic requirements of the traffic across the provisioned Network Slice, thus, paving the way for autonomous and self-optimized 5G-ready optical networks.

## ACKNOWLEDGEMENT

This work has been supported by the H2020 5GPPP SLICENET project (H2020-ICT-2016-2/761913) and the Spanish Government through project ALLIANCE-B (TEC2017-90034-C2-2-R) with FEDER contribution.

## REFERENCES

- [1] 5G-PPP 5G Architecture White Paper, Version 2.0, December 2017.
- [2] NGMN Alliance, "Description of network slicing concept", January 2016.
- [3] ONF TR-526, Applying SDN Architecture to Network Slicing, Issue 1, April 2016.
- [4] R. Vilalta et al. "Multitenant Transport Networks With SDN/NFV," *IEEE Journal of Lightwave Technology*, vol. 2, no. 6, pp. 1509–1515, March 2016.
- [5] J. Baranda et al., "Orchestration of End-to-End Network Services in the 5G-Crosshaul Multi-Domain Multi-Technology Transport Network", *IEEE Communications Magazine*, vol. 56, no. 7, pp. 184–191, July 2018.
- [6] 3GPP TR 28.801, Study on management and orchestration of network slicing for next generation network, Version 15.1.0, January 2018.
- [7] IETF draft-ietf-supra-generic-policy-info-model-03: "Generic Policy Information Model for Simplified Use of Policy Abstractions (SUPA)", May 30, 2017.
- [8] ETSI GS NFV-MAN 001, Network Functions Virtualization (NFV); Management and Orchestration, Version 1.1.1, December 2014.