INTERNET TECHNOLOGY LETTERS

Hierarchical Service Providers on Sliced Infrastructure

Journal:	*Internet Technology Letters*
Manuscript ID	ITL-19-0008.R1
Wiley - Manuscript type:	Special issue article
Date Submitted by the Author:	02-Apr-2019
Complete List of Authors:	Tusa, Francesco; University College London, Electronic and Electrical Engineering Clayman, Stuart; University College London, Electronic and Electrical Engineering Valocchi, Dario; University College London, Electronic and Electrical Engineering Galis, Alex; University College London , Electronic and Electrical Engineering
Keywords:	Slicing, MANO, VNF, VIM, VIM on-demand

Received: Added at production | Revised: Added at production | Accepted: Added at production

DOI: xxx/xxxx

ARTICLE TYPE

Hierarchical Service Providers on Sliced Infrastructure

Francesco Tusa* | Stuart Clayman | Dario Valocchi | Alex Galis

[1]Dept. of Electronic and Electrical Engineering, University College London, London, United Kingdom

Correspondence
*Email: francesco.tusa@ucl.ac.uk

Present Address
Torrington Place, London WC1E 7JE, UK

Summary

This work presents the deployment of a hierarchical multi-MANO environment for the instantiation and management of end-to-end 5G network services in a slice-enabled infrastructure.

KEYWORDS:
Slicing, MANO, VNF, VIM, VIM on-demand

1 | INTRODUCTION

In the 5G landscape, different Infrastructure and Service Providers can explore new business models by federating their resources and service offerings and providing their customers with the capability of instantiating end-to-end services across multiple technological resource domains. The service components can be located in separate geographical locations of one Provider or even outside of its own administrative borders. In order to facilitate the allocation of resources for the end-to-end services, each 5G Provider may take advantage from using complex software systems that perform the management, control and orchestration of the different resources that need to be allocated for the deployment of 5G Service Function Chains. Each of these systems, defined by ETSI [†], is usually referred to as MANO (***Man***agement and ***O***rchestration of VNF) as it is able to deal with the tasks required to build Network services by combining multiple Virtual Network Functions (VNFs).

In this paper, we consider a 5G Provider's resource infrastructure consisting of two or more geographically separate offices (or business units) that seek to deploy network services composed of particular types of interconnected VNFs, whose elements location could potentially be anywhere in the distributed resource infrastructure – from the mobile edge to the core. Furthermore, the physical resources utilized in each segment of the infrastructure can be managed by a particular MANO system that a business unit of that Provider is not willing to replace or update. Different MANO systems operating on dedicated segments of the distributed NFVI (for NFV Infrastructure) can inter-work according to a hierarchical topology, using north-south interfaces, in order to support the instantiation of more complex end-to-end services. More specifically, each south-bound interface of a MANO system will interact with the north-bound interface of its counterpart, just as it would do with a VIM (Virtual Infrastructure Manager). This interaction will not have to be limited to just a single layer (i.e., one upper-MANO and one lower-MANO) and might further be extended to a multi-layered stacked scenario where a recursive approach could simplify both the operational management and the creation of new services – the most general case of this recursive architecture would be a tree – the Hierarchical Service Provider (HSP). To support the concept of *slicing*, we present an extension to the above HSP multi-MANO north-south interaction, with an implementation of a scenario based on existing MANO systems that have been devised and already utilized in the context of 5G. In order to support service provisioning over a *slice-enabled* distributed NFVI, mechanisms that enable the slicing of the whole end-to-end infrastructure – from the mobile edge to the core Data Center (DC) – including network, compute and storage resources are required.

Slicing is a move towards the segmentation of resources and the deployment of NFV for the purpose of enhanced services and applications on a global shared infrastructure. To manifest this slice approach, we have designed and built a *DC Slice Controller* that is able to allocate a slice of a DC and create a per-slice VIM (Virtual Infrastructure Manager) in an on-demand fashion. As

[†] ETSI . ETSI GS NFV-MAN 001 V1.1.1. Network Functions Virtualization (NFV); Management and Orchestration. https://www.etsi.org/deliver/etsi_gs/NFV-MAN/001_099/001/01.01.01_60/gs_NFV-MAN001v010101p.pdf.

each slice and its VIM instance are independent of the others, improved resource isolation will be provided to the customers, who will never share the same physical and virtual infrastructures. Each of these slices will be allocated and de-allocated on-demand, via interacting with the DC Slice Controller. As each slice will be separate from the others and will include its own VIM instance, a more effective resource management will be ensured thanks to the isolation introduced at both the control and data planes. Moreover, having different types of VIMs in different segments of the infrastructure will allow flexibility creating end-to-end slices according to the type of resources that will host that particular VIM: lightweight VIM versions might be instantiated at the mobile edge, whereas the core DC might host more traditional VIM implementations, such as OpenStack.

The reminder of this paper is organised as follows: Section 2 reports about the current state-of-the-art on slicing; Section 3 presents relevant slicing concepts and approaches; Section 4 describes the architectural elements, interactions and initial relevant results for the considered multi-MANO scenario; Section 5 draws some conclusions and proposes possible future works.

2 | RELATED WORK

The desire to provide services on top of slices, by logically partitioning resources of a multi-domain software-defined infrastructure can be obtained through different slicing strategies depending on which system elements provide the slicing and at what layer the slicing is introduced. The overall concept of network slicing is outlined by various standards organisations including ETSI [‡], the IETF [§], and NGNM [¶]. If slicing is co-ordinated by the Orchestrator, which uses an inter-domain orchestrator API interaction and / or a peer to peer approach, a slice is closer to a small abstraction in the Orchestrator, rather than an infrastructure partition. There are various projects and initiatives that are doing slicing at the Orchestrator level. These include 5G-TRANSFORMER [#], 5G!PAGODA [‖], SliceNet [**]. Their approach has the easiest entry position, but all the main software elements need to be updated and adjusted to know about slices; and all of the APIs, the modules, and the data structures need to be adjusted and adapted to factor in slices. Consequently, there are inherent trade-offs when selecting one or the other slicing approach. The actual decision on which slicing approach is realized depends on various key aspects of the service requirements under consideration, and can be focused on the technical requirements of the provider, and the technological choices of the tenants. The recent survey [1] has been looking at approaches to 5G slicing. Additional characteristics, standard and research activities on slicing and references can be found here [2].

Slicing at lower layers, namely infrastructure slicing, means that upper layers, such as VIMs and Orchestrators do not need to directly know about slicing. If a slice is presented to one of these VIMs or Orchestrators, they can carry on working with no change or minimal change. Such an approach has been presented in Clayman [3] and Freitas [4]. Initial results of Slicing, while using a hierarchy of interacting MANOs, are presented in Tusa [5]. The NECOS project [††] also uses an alternative approach of infrastructure slicing to manifest a Slice as a Service functionality. It uses a single orchestrator, with a multi-domain view, and a set of on-demand VIMs to manage sliced resources which encompass mobile edge and core elements. A combination of Slice Providers, an online Marketplace, and Slice Controllers are used to dynamically build an end-to-end slice for service deployment.

3 | SLICING CONCEPTS

While working on the design and implementation of the multi-MANO scenario, we were faced with a fundamental question about slicing: how is it possible to connect the various resource segments (including the mobile edge, the core networks, and the data centers) and which layer of the existing MANO architecture should implement the slicing. In particular, we identified two possible answers to this question: namely, either (i) changing the existing orchestration and management software elements to deal with slices, or (ii) considering slicing at a lower layer and presenting a slice to the upper layers. If we choose the former approach, then all the main software elements would need to be updated and adjusted to know about slices, together with all

‡ ETSI Report on Net Slicing Support with ETSI NFV Architecture Framework. http://www.etsi.org/deliver/etsi_gr/ NFV-EVE/001_099/012/03.01.01_60/gr_NFV-EVE012v030101p.pdf

§ IETF. Network Slicing - Revised Problem Statement, draft-galis-netslices-revised-problem-statement-03. 2018.

¶ NGNM. Description of Network Slicing Concept, NGMN 5G P1 Requirements & Architecture, Work Stream End- to-End Architecture. 2016. https://www.ngmn.org/fileadmin/ngmn/content/images/news/ngmn_news/NGMN_5G_White_Paper_V1_0.pdf

5G-TRANSFORMER. 5G Mobile Transport Platform for Verticals. 2017. http://5g-transformer.eu

‖ 5G!PAGODA. EU-Japan - A network slice for every service. 2017. https://5g-pagoda.aalto.fi

** SliceNet. End-to-End Cognitive Network Slicing and Slice Management Framework in Virtualised Multi-Domain, Multi-Tenant 5G Networks. 2017. https://slicenet.eu

†† NECOS. Novel Enablers for Cloud Slicing.www.h2020-necos.eu/

their related APIs. We realized that if slicing is instead implemented at a lower layer, then the upper layers, such as VIMs and Orchestrators, would not need to know anything about slicing[3]. The following sections describe the slicing model based on the second approach, which makes use of the new concept of Data Center slicing and VIM on-demand. We present how a Data Center can be sliced and how a VIM can be allocated for the slice on-demand.

VIM on-Demand Slices

A Slice is an aggregated set of resources that can be used in the context of an end-to-end networked service comprised of VNFs. Slicing allows composing multiple resources into logically isolated network partitions, with a slice being considered the basic unit of programmability that includes network, computation and storage. When analyzing the wide variety of applications to be supported by 5G networks, it is necessary extending the concept of slicing to cover a wider range of use-cases than those targeted by the current SDN/NFV technologies. If we have slicing everywhere, including networks and DCs, we observe the following: (i) there is a separation of physical resources; (ii) there is isolation of services as no customers share physical resources; and (iii) it is secure as only a given specified customer can access a host, so that there will be no sharing or cross VM issues.

In order to support service provisioning over these slices, it will be necessary having mechanisms that implement the slicing of both the network resources and the Data Center compute and storage resources across the different segments of a Provider's infrastructure. To manifest this slice approach, we have designed and built a DC Slice Controller that is able to allocate a slice of a DC and create a per-slice VIM in an on-demand fashion. The DC slice and the VIM are provisioned solely for use with the service. Each slice and its associated VIM are independent of the other slices and VIMs. In this way, customers will never share servers, and the worry of VMs of one customer interacting or spying on another customer will be eliminated. The issue of one customer's VM consuming all the resources and starving other customer's VMs is also ameliorated to some extent.

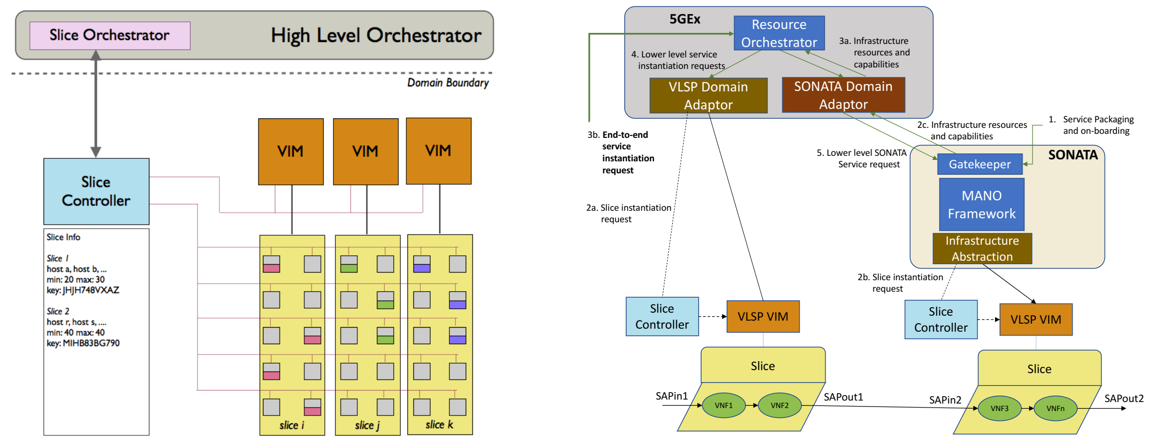
Each of these slices will be allocated and de-allocated in an on-demand fashion, under software control via interacting with the DC Slice Controller. Moreover, the resource management of the whole infrastructure will not be based on a shared model, as independent VIM instances will be allocated per slice, and the best type of VIM can be selected for deployment according to the type of resources available in each part of the segmented end-to-end infrastructure. Figure 1a presents how the resources of a DC are isolated from each other, and how a Slice Controller is involved in such a process.

A **DC Slice Controller** supports Data Center slicing and the VIM on-demand model. The main components and functions of the Slice Controller use the following elements for its operation: (i) a *Resource Manager* – which manages all of the resources in the DC and keeps a track of which resources have been allocated to which slice. (ii) a *Slice Information Store* – which database lists all of the slices and all of the resources in the slice. (iii) a *Slice Creator* – which is responsible for handling requests for slices and interacting with the Resource Manager to determine if it is possible to create a new slice. If the slice creation is possible, it interacts with the VIM Factory. (iv) the *VIM Factory* – which is able to allocate a VIM and configure it to use the resources which have been picked by the Slice Creator. Once the VIM is allocated and deployed, the REST entry point is returned. (v) the *VIM Placement Manager* – which is responsible for determining which host should be used to execute a newly created VIM.

4 | END-TO-END MULTI-MANO HSP SYSTEM

Different MANO architectural elements can be combined in order to build a hierarchical slice-enabled multi-MANO environment for the deployment and management of end-to-end 5G network services. The architectural view of the multi-MANO interaction is depicted in Figure 1b. The implementation uses two different MANOs, with a 5GEx framework instance[6] at the top level of the hierarchy, playing the role of the higher level orchestrator. This was chosen as the 5GEx system was originally conceived and built to be a multi-domain orchestrator. At the lower level of the orchestration hierarchy, an instance of the SONATA Service Platform[7] is deployed to operate on the resources of its own segment of the NFVI, and behaving as a local Domain Orchestrator, connected to the top 5GEx MANO via the SONATA Domain Adapter.

The 5GEx MANO operates over its own segment of the NVFI using the DC Slicing approach discussed in Section 3. The on-demand allocation of slices containing lightweight VLSP (Very Lightweight Network & Service Platform[8]) VIM instances, is triggered by the VLSP Domain Adapter interacting with the local Slice Controller. The VLSP Domain Adapter plays the same role played by the Slice Orchestrator component in Figure 1a. In a similar way, the SONATA MANO operates on top of another segment of the NFVI, again using the DC Slicing capabilities, and the slices include separate on-demand instances of the VLSP VIM. SONATA's own Infrastructure Adapter drives the creation of those slices by communicating with the local Slice Controller and offers to the SONATA MANO the interface to interact with the on-demand VIM allocated in each slice. The key point of this specific multi-MANO inter-working solution is in the interplay between the SONATA Domain Adapter in the 5GEx MANO



(a) DC Slicing via the Slice Controller **(b)** Step-by-step multi-MANO interaction using VLSP slices

FIGURE 1 Architectural elements of the considered multi-MANO scenario and their interaction

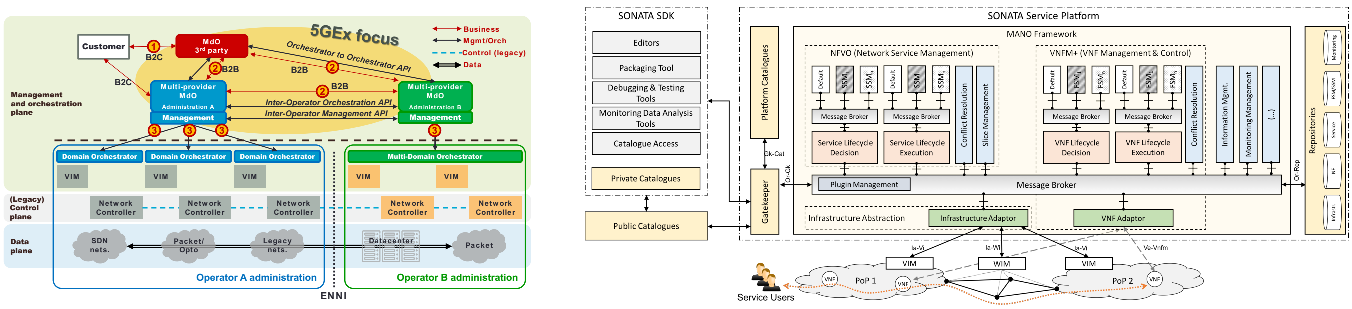
(left side of Figure 1b) and the Gatekeeper in the SONATA MANO (right side of Figure 1b). The SONATA Domain Adapter utilizes the REST API exposed by the Gatekeeper, and provides the 5GEx Resource Orchestrator with an abstracted view of the resources and services available from the slices of the underlying SONATA domain. More specifically, the services and the VNFs available from the SONATA MANO are exposed by the SONATA Domain Adapter to the 5GEx Resource Orchestrator as Domain Capabilities, together with an abstracted view of the resources available from the allocated slice. When a request for instantiating an end-to-end service is received at the top-layer MANO, the 5GEx Resource Orchestrator therein will deploy the requested service elements in the different available DC slice segments according to the resources and capabilities reported by the corresponding Domain Adapters. Virtual links will be created between the allocated service elements (e.g., VNF1, VNF2 and VNF3 in Figure 1b) for the implementation of the requested end-to-end Service Function Chain.

4.1 | Multi MANO HSP Architecture Elements

This section further describes the architectural elements utilized in the context of a 5G-PPP cross-projects activity aimed at the realization of a multi-MANO Hierarchical Service Provider (HSP) scenario leveraging on the systems developed by the 5GEx [6] and SONATA [7] EU projects. Relevant data models, abstractions and interfaces from the above projects, which are relevant for the implementation of the inter-working multi-MANO system, are presented to highlight how end-to-end network services can be deployed on multi-domain end-to-end slices using a 5GEx MANO and one or more instances of a SONATA MANO.

5GEx: The 5GEx project architecture framework [6] depicted in Figure 2a identifies the main components and the inter-working interfaces involved in the multi-domain orchestration. The top of the figure shows the key 5GEx MANO component (i.e., the Multi-domain Orchestrator – MdO); the middle part shows the Domain Orchestrators, responsible for performing resource orchestration on the actual infrastructure. The MdO handles the orchestration of resources and services from different providers, coordinating resources and services orchestration at the multi-domain level (either multi-technology or multi-provider). The 5GEx I3 interface (from Management to Domain Orchestrator) was designed and implemented to allow the interaction between a MdO and the underlying Domain Orchestrators. In particular, the sub-interface I3-RC is responsible for retrieving the local resource availability from the technological Domains and for enabling their programmability. This interface provides the abstractions required for the implementation of the multi-MANO north-south interaction.

SONATA: The SONATA project [7] is based on a customizable Service Platform with a NFV Orchestrator that supports Network Service Software Development Kit (SDK) for developers and specialized DevOps workflows. SONATA's MANO components are represented on Figure 2b. The Gatekeeper module is the main entry point of the Service Platform and implements its Northbound Interface. It exposes API endpoints for SDK tools and provides the following main functionalities: (i) accepting new services, in the package format, to be deployed on the platform; (ii) accepting and orchestrating new service requests from customers interested in instantiating a service; and (iii) following a service performance through automatically monitoring each on-boarded service or function. The Gatekeeper plays an important and active role in the implementation of the communication



(a) 5GEx Framework

(b) SONATA Framework

FIGURE 2 Systems of the considered multi-MANO scenario: **(a)** 5GEx [6] and **(b)** SONATA [7] frameworks

interface in our multi-MANO HSP scenario, as it provides the API methods for the interaction between the 5GEx MANO and SONATA MANO. Further details are reported in the following section where we show how all these elements are composed to form the full hierarchy.

4.2 | Experimental Implementation and Evaluation

For the validation of the devised multi-MANO scenario, we consider a service developer who wishes to deploy a service on a 5G Provider infrastructure, where one branch of the provider is managed by the 5GEx MANO and another one uses the SONATA (SP) MANO system. As a prerequisite, it is important that all the required service elements and VNFs are on-boarded before any deployment can occur. For 5GEx, a standard set of VNFs and their descriptors are on-boarded when the MANO starts, and services get composed from those VNFs; however, on the SONATA MANO, a packaged service has to explicitly be on-boarded beforehand via the Gatekeeper. On initialization, in the branch that makes use of the 5GEx MANO system, the VLSP Domain Adaptor creates an on-demand light-weight VLSP DC slice and stores its relevant parameters (Step 2a in Figure 1b). Analogously, in the Provider's branch using the SONATA MANO, the slice creation process is driven by the Infrastructure Abstraction (IA) by interacting with the local Slice Controller (Step 2b on on Figure 1b). Another on-demand instance of the lightweight VLSP VIM is allocated as part of another DC slice, and references to it are returned back to the IA.

When the DC slices creation process in each NFVI segment is completed, the 5GEx Resource Orchestrator will be able to build the overall end-to-end slice view by combining together the DC slices reported by the lower level adaptors. A key-role in the creation of the end-to-end slice view is played by the SONATA Domain Adaptor (in the 5GEx MANO), which implements the abstraction and mechanisms required to enable the inter-working of the two above MANOs. At the southbound, the SONATA Domain Adaptor interacts with the SONATA Gatekeeper (Step 2c on Figure 1b) to gather information about resources, services, and functions that are available from the allocated slice. The low-level details will be transparent to the 5GEx Resource Orchestrator, which can only collect an abstracted view from each domain via the I3-RC interface (Step 3a on Figure 1b).

An end-to-end service instantiation request can now be issued to the 5GEx Resource Orchestrator, which will split the required end-to-end service elements on the lower-layer resource domains according to the capabilities and resource availability reported by each of them (Step 3b on Figure 1b). Next, embedding occurs where lower level service instantiation requests are sent to the individual Domain Adapters according to the embedding decisions of the 5GEx MANO's Resource Orchestrator (step 4 on Figure 1b). The end-to-end service request includes service elements available from the SONATA domain, and so the SONATA platform will receive a service instantiation request from the SONATA Domain Adapter on the relevant endpoint of the Gatekeeper API (step 5 on Figure 1b). From the SONATA SP perspective, the request will look like a normal service request, such that the SONATA service deployment flow will follow the request, instantiating the entities as described in the package on-boarded by the service developer in Step 1.

We evaluated this setup on the UCL (University College London) testbed implementing a test service as an orchestrated deployment of VLSP lightweight VNFs across the sliced NVFI. After successful instantiation of the requested end-to-end service, network traffic was generated and injected in the Service Function Chain (which only included forwarding VNFs) using the iperf network bandwidth measurement tool. More specifically, an iperf server was attached to the SAPout2 of Figure 1b and was able to receive the packets generated by an iperf client instance attached to SAPin1.

5 | CONCLUSIONS

This paper considered emerging 5G scenarios where different segments of a Service Provider's infrastructure – from the edge cloud to the central DC – are administered by separate organisational divisions / departments, each possibly relying on a particular MANO systems for the deployment of the end-to-end services on the distributed resource infrastructure. Whilst each domain has its own MANO, the MANO instances are configured in a north-south way creating a hierarchy of service provision capabilities, called the Hierarchical Service Provider (HSP). This approach works particularly well where each domain, from the mobile edge, to the core DC, can be managed independently of the others, but needs to be combined to form slices.

We discussed how the end-to-end slicing approach can provide a more effective resource management for the deployment of customer services. We demonstrated why having a multi-MANO hierarchical orchestration system with slicing support can be a viable approach to address the requirements of that scenario, and we presented this unique and new configuration where services are deployed across distributed domains. In this paper, we described a design and implementation to the problem, and provided a working solution to the HSP scenario, that includes both the 5GEx and SONATA MANO frameworks. Whilst our HSP experiment is still exploratory work, on the very wide subject of MANO platforms inter-working, we demonstrated how the proper combined use of data models, abstractions, and APIs provided by those systems could be considered as a reference for future work on the design and implementation of more generic HSP multi-MANO scenarios.

As further work, we plan also to evaluate the same infrastructure against more common or commercial VIMs, such as OpenStack or Xen with the particular focus on viable deployment options, economy of scale in slicing, support for service diversity, and the interaction with vertical customers/tenants. We plan to design and further evaluate efficient mechanisms for the inter-connection of the different slice parts, for slice stitching, and for slice management enablers in multi-MANO scenarios.

References

1. Afolabi I, Taleb T, Samdanis K, Ksentini A, Flinck H. Network Slicing and Softwarization: A Survey on Principles, Enabling Technologies, and Solutions. *IEEE Communications Surveys and Tutorials* 2018; 20(3): 2429–2453.
2. Galis A. Net Slicing – Tutorial. In: IEEE CNSM 2018. ; 2018; Rome, Italy. http://cnsm-conf.org/2018/tutorials.html.
3. Clayman S, Tusa F, Galis A. Extending Slices into Data Centers: the VIM on-demand model. In: IEEE 9th International Conference on Network of the Future - NoF 2018. ; 2018; Poznań, Poland.
4. Freitas LA, Braga VG, Correa SL, et al. Slicing and Allocation of Transformable Resources for the Deployment of Multiple Virtualized Infrastructure Managers (VIMs). In: S4SI Workshop - Netsoft. ; 2018; Montreal, Canada.
5. Tusa F, Clayman S, Valocchi D, A.Galis . A Multi-Domain Orchestration for the Deployment and Management of Services on a Slice Enabled NFVI. In: MOBSLICE, NFV-SDN. ; 2018; Verona, Italy.
6. Vaishnavi I, Czentye J, Gharbaoui M, et al. Realizing services and slices across multiple operator domains. In: NOMS 2018 - 2018 IEEE/IFIP Network Operations and Management Symposium. ; 2018: 1-7.
7. Dräxler S, Karl H, Peuster M, et al. SONATA: Service programming and orchestration for virtualized software networks. In: 2017 IEEE International Conference on Communications Workshops (ICC Workshops). ; 2017: 973-978.
8. Clayman S, Mamatas L, Galis A. Efficient Management Solutions for Software-Defined Infrastructures. In: 5GMan - IFIP/IEEE Network Operations and Management Symposium (NOMS). IEEE/IFIP. ; 2016.

Acknowledgements

This work was supported by the EU project NECOS – "Novel Enablers for Cloud Slicing" (777067).