Proposal and Investigation of a Scalable NFV Orchestrator Based on Segment Routing Data/Control Plane

V. Eramo(*), F.G. Lavacca(**), T. Catena(*), M. Polverini(*), A. Cianfrani(*)

(*)DIET, "Sapienza" University of Rome - Via Eudossiana 18, 00184 Rome, Italy
(**)Fondazione Ugo Bordoni, Viale del Policlinico 147, 00161 Roma, Italy
{*vincenzo.eramo,francescogiacinto.lavacca,marco.polverini,antonio.cianfrani*}*@uniroma1.it*
{*fglavacca*}*@fub.it*

***Abstract*—Network Function Virtualization (NFV) has attracted significant interest in recent years due to its agility to adopt new technology and its flexibility in on-demand deployment of network functions. A major topic in the field of NFV research is the allocation and orchestration of VNF instances to provide satisfying network services according to user demands. The paper proposes and investigates a new solution for a scalable NFV Orchestrator allowing for a reduction of the needed time for the Service Function Chain routing and the allocation of cloud and bandwidth resources. That is obtained by abstracting and simplifying in the NFV orchestrator the knowledge of both the network topology and the resource amount available in the datacenters in which the Virtual Network Functions are executed. The solution is based on the Segment Routing paradigm that allows to implement the scalable orchestration at data plane level. Finally we show how the proposed solution reduce by 90% the time needed to execute an SFC routing and resource allocation algorithm with respect to a traditional orchestration solution.**

***Index Terms*—Network Function Virtualization, Scalable Orchestrator, Segment Routing, Software Defined Network.**

I. INTRODUCTION

The Network Function Virtualization (NFV) concept is one of the most appealing research outcomes for network operators. The research activity on NFV has been focused on the solution of the SFC Routing and Cloud and Bandwidth resource Allocation (SRCBA) problem [1]; it is composed of two subproblems: i) the Service Function Chaining (SFC) routing and ii) the allocation of network and computing resources for the execution of Virtual Network Function Instances (VNFI). The performance indexes used to characterize an SRCA algorithm are the number of SFCs accepted, the server power consumption [2], [3], the bandwidth used to interconnect the VNFIs [4], the cloud resource amount needed to activate the VNFIs and so on [5], [6], [7].

All of these solutions assume that the NFV orchestrator, responsible to implement the SRCBA algorithm, knows in detail the network topology and the available resources in the datacenter in which the VNFIs have to be instantiated. That leads to a high computational complexity of the SRCBA algorithm increasing versus the number of servers and switches located in the datacenters [8].

Furthermore this traditional solution may not be applicable in multi-provider NFV environments that implement business models with the following two players [7],[9]-[11]: the Cloud Infrastructure Provider (CInP) and the Internet Service Provider (ISP). The function of an CInP is to deploy and manage the physical resources on which the virtual resources (i.e. Virtual Machine) may be provisioned. The function of an ISP is to rent the physical resources from one or more than one CInP to execute the VNFIs. It provides to interconnecting VNFs through bandwidth resources to create services to the end users. It also manages the NFV Orchestrator that acquires from the CInP through the Virtual Infrastructure Manager (VIM), information about the network topology and the available resources (bandwidth, processing, disk memory) of the Cloud Infrastructure. When CInP and ISP are two different players, the CInP could need to hire from the ISP some of these information and only a partial description could be sent to the NFVO. For this reason the application of classical SRCBA algorithms is not possible in a Multi-Provider NFV environment.

In this paper we propose and investigate a scalable solution for the resource orchestration in NFV networks. In the proposed solution the VIM, responsible for the management of a single datacenter, only notifies the NFVO information about the amount of available cloud resources without providing any detail on the network topology of the datacenter and the available number of servers. One of the main feature of our solution is to exploit Segment Routing (SR) paradigm [12], [13], [14] to implement the data plane of SFC routing and to hide the complexity of datacenters infrastructures.

The main innovative contributions of the paper are the following:

- the proposal of a scalable solution for the resource orchestration in Multi-Provider NFV environments;
- the definition of a SR-based data/control plane for the SFC routing applicable for the proposed scalable orches-

978-3-903176-14-0 © 2018 IFIP

tration;

- the investigation of computational complexity reduction with respect to a traditional orchestration reduction.

We illustrate the scalable orchestration solution in Section II. The advantages in terms of reduction in computational complexity are evaluated in Section III. Finally conclusions and future research items are described in Section IV.

II. The Scalable NFV Orchestrator

The network scenario considered is the one of an Internet Service Provider (ISP), composed by a core network interconnected to multiple datacenters. The datacenters, hosting computing resources for the execution of VNFs, can be managed by the ISP provider or by different CInPs. The orchestration architecture is based on the ETSI standard [15] described in Fig. 1. It details the key functional blocks in NFV system: Virtualized Network Function (VNF), NFV Infrastructure (NFVI) and NFV Management and Orchestration (MANO) framework [15]. The MANO comprises of the Virtualized Infrastructure Manager (VIM), VNF Manager (VNFM) and NFV Orchestrator (NFVO). The communication between various functional blocks in NFV architectural framework is enabled through a set of well-defined reference points. The VNF

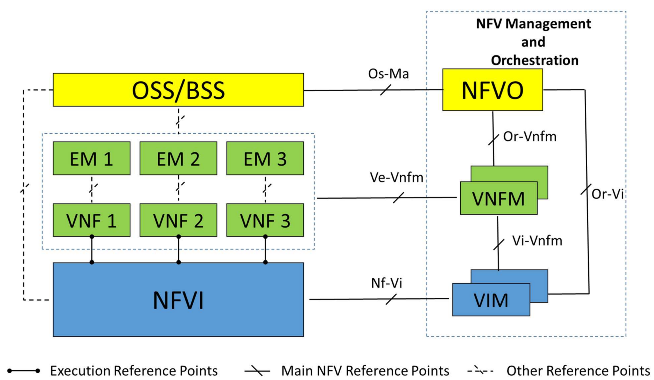

Figure 1. ETSI NFV Standard.

Instance is the software implementation of a network function. It runs on NFV Infrastructure (NFVI) that encompasses the necessary software and hardware components. NFVI can span geographically distributed locations, called NFVI Point of Presences (NFVI-PoPs). NFVI resources (e.g. compute, storage and network) can be managed and controlled by one or more VIMs. The VIM is in charge to instantiate and manage the cloud resources (processor cores, RAM memory, disk memory) used by the VNF [16]-[17]. The VNFM is responsible for the lifecycle management (e.g. instantiation, scaling, healing and monitoring [18][19]) of one or more VNF instances. The NFVO is in charge of the orchestration of NFVI resources across multiple VIMs and lifecycle management of network services. The NFVO and VNFM work jointly to ensure that the network services and their corresponding VNFs meet the service quality requirements. Thus, the performance of NFVO and VNFM is crucial to NFV system.

In the traditional operation mode, the NFVO 1 acquires from the VIMs, information about the network topology and available resources as illustrated in Fig. 2.a where the NFVI-POP-1 sends to the NFVO the information about a network topology composed by 4 servers and three switches connected according to a tree topology; a number of two processing cores per each server is also notified the NFVO. There are the following problems in the implementation of this solution:

- the NFVO needs to acquire many informations and the graph representing the entire network topology and the available cloud and bandwidth resources can become very complex;
- the SFC Routing and Cloud and Bandwidth resource Allocation (SRCBA) algorithm does not scale with the increase in both number of servers and switches of the NFVI-PoP networks;
- the solution is not applicable in the case of multi-provider NFV environments in which the providers of each NFVI-PoP may need to hire the organization of own NFVI-PoP network and not to make it visible to the owner of the NFVO.

Our idea is i) to associate a VIM to each NFVI-PoP so that to reduce the SRCBA algorithm complexity, and ii) to use Segment Routing to implement data plane forwarding. More in detail, we propose the scalable solution depicted in Fig. 2.b, where the VIM only notifies the NFVO of about the available cloud resources; the servers status and the NFVI-PoP network topology is not reported to the NFCO; for instance the VIM only notifies the NFVO that 8 processing cores are available and can be allocated to the VNFI executing the Service Functions (SF) of any SFC. This solution drastically decreases the complexity of the SRCBA algorithm. Its output will be the NFVI-PoP hosting each VNFIs and the amount of cloud resources to be assigned, without specifying the routing paths followed by the SFC inside the NFVI-PoP networks. The VIM will be responsible for internal NFVI-PoP paths computation, by trying to balance as much as possible the load of the VNFIs executing a given Service Function according to the number of processing cores assigned to the VNFIs.

A. Segment Routing Based SFC Routing

For the data plane we propose the use of SRv6, i.e. Segment Routing over IPv6. SRv6 allows to code SFC routing in a very easy and flexible way. The SRv6 paradigm is based on explicit and source routing features: the ingress node of a SR network is able to code the network path to be followed by a packet inserting into the SR header a Segment List, i.e. a list of Segment Identifier (SID). The Segment List is a list of nodes to be crossed by the packet; the final path will be composed by the IPv6 paths interconnecting the nodes reported in the Segment List.

In this work we exploit a feature of SRv6: the possibility of associating several identifiers, referred to as SIDs, to the same node. In this way, a specific SF running in a NFV-PoP can be identified by a SID associated to the edge node connecting the NFV-PoP to the core network, independently of the specific

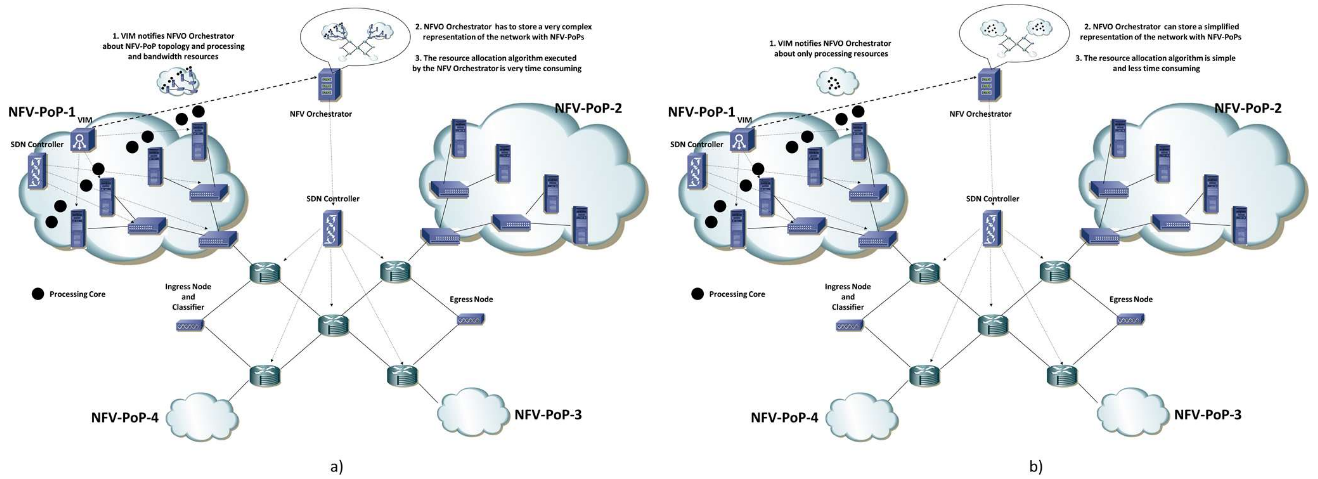

Figure 2. Traditional (a) and Scalable (b) NFV Orchestration.

instance (NFVI) where the SF is running. In other words, SRv6 allows to abstract the datacenter resources at data plane level. To better explain this concept, let us consider the example reported in Fig. 3 where a seven nodes core network with two NFV-PoPs is providing SFC composed by two functions, a Firewall (FW) and an Intrusion Detection System (IDS). Each core node have its own SID: for the easieness of representation we associate to each node a SID having as first bits the node name and as last bits 1, i.e. A SID is A::1. The FW and IDS SFs running at NFV-PoP-1 are identified by two SIDs of router A, i.e. A::F1 and A::F2 respectively; in a similar way, D::F1 and D::F2 are the SID of FW and IDS SFs running at NFV-PoP-2. When router A receives a packet having A::F1 as active segment, it send the packet towards the local NFV-PoP where a dedicated mechanism will be responsible for the delivery to the proper VNFI. In our proposal we consider an SDN-based solution for the NFV-PoP forwarding, as proposed in [20], that allows to balance the load towards the set of VNFIs that execute the SF.

In Fig. 3 we show the overall set of operations performed for a packet crossing the ISP network. In the example we consider router E1 as the Ingress router, i.e. the one responsible for packet classification and SR-based SFC routing (FW + IDS), and router E2 as Egress node. The Segment List used to encode the SFC is $\{$ A::F1, D::F2, E2::1$\}$, where A::F1 identifies the FW SF running at NFV-PoP-1, D::F2 identifies the IDS running at NFV-PoP-2, and E2::1 identifies the engress node E. The FW and IDS SFs are executed in VNFIs instantiated in the NFVI-PoP-1 and NFVI-PoP-2 respectively. Four VNFIs are instantiated for FW SF in the servers S1, S2, S3 and S4 respectively. One processing core is allocated in servers S1, S2 and S3 while two cores are allocated in server S4. An SDN controller provides to configure the access switch S1 and the other switches S2 and S3 in the NFVI-PoP-1 in order to implement a load-balancing solution towards the activated VNFIs. In particular we adopt the Balanced Hash Tree (BHT) load-balancing solution proposed in [20]. BHT implements load balancing on the switches through the select group table, as described in the OpenFlow specification, which requires processing to be based on a switch-computed selection algorithm. In BHT, switches use a hashing-based algorithm to determine the output port for load balancing towards VNFIs and the switch uses the select group type to assign each egress flow to an action bucket. Flow distribution is similar to the well-known equal cost multipath routing (ECMP) strategy, which balances loads using multiple equal-cost paths between two neighboring hops. However, BHT is different because it also considers service chaining between SFs and can assign weights to different paths. This is illustrated in Fig. 3 where for the switch SW1 the weigths of the select entry is chosen equal to 2 and 3 for the O1 and O2 output interfaces of the switch. In such a way $3/5$ and $2/5$ of the packet flow will be directed toward O1 and O2 in order to take into account that these interfaces allow to reach VNFIs having two and three processing cores allocated, respectively.

III. NUMERICAL RESULTS

We provide some results to compare the computational complexity of the Traditional Orchestration (TO) and proposed Scalable Orchestration (SO) solutions. In particular we evaluate for the two solutions the execution times of the SFC Routing and SFC Routing and Cloud and Bandwidth resource Allocation (SRCBA) algorithm proposed in [21] and applied for the two solutions. In both cases a multistage graph is built and a shortest path is computed to decide the cloud and bandwidth resources to be used. The graph is composed by a number of nodes equal to the number of servers and datacenters in the case of traditional and scalable NFV orchestrator respectively. The fewer nodes in the proposed solution leads to lower computation times for the scalable case. The times are measured on a Personal Computer characterized by 3.40 GHz Intel i7-3770 processor and by an 8 GB RAM memory.

The comparison is carried out when a realistic scenario is

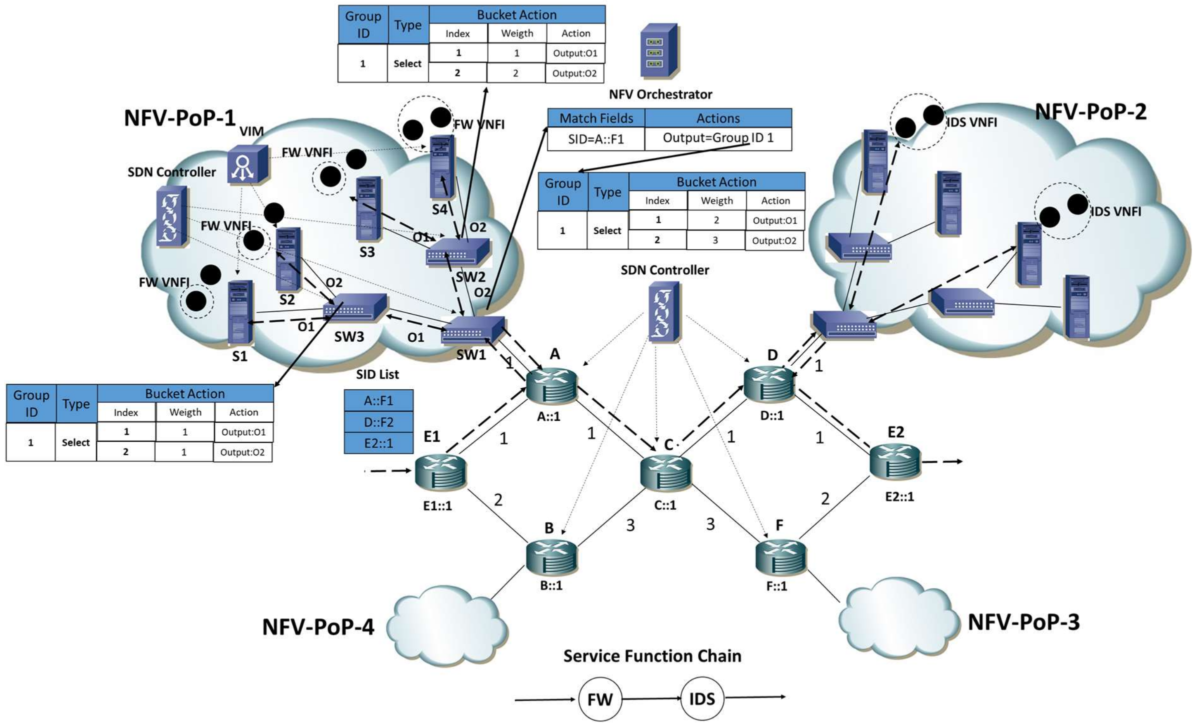

Figure 3. Segment Routing based Data/Control Plane for the scalable NFV orchestration architecture.

considered. The Deutsche Telekom medium distance network reported in Fig. 4 with 13 nodes is considered. One NFVI-PoP is located only in the nodes of Frankfurt, Leipzig, Nuremberg and Hannover. Each NFVI-PoP is equipped with 4 servers connected to the access switch with a fat tree network. We assume that the NFVI-PoPs are equipped with 96 cores, that is 24 per each server. The link bandwidth of the network interconnecting the NFVI-PoPs is fixed equal to 40Gb/s.

We consider four Service Functions: Firewall (FW), Intrusion Detection System (IDS), Network Address Translator (NAT) and Proxy. VNF instances can be instantiated in the NFVI-PoPs to support these SFs. They are supported by software modules characterized by the maximum processing capacity and by the number of cores reported in Table I. The possible compositions of the SFs in SFC are reported in Fig. 5 where the nodes v_s and v_e are the access nodes originating and terminating the SFC. We assume that N SFC requests are generated. The bandwidth values of the SFCs are chosen equal to 20 Mbps. The ingress and egress nodes of each SFC is randomly chosen.

Table I
MAXIMUM PROCESSING CAPACITY AND ALLOCATED NUMBER OF CORES FOR THE SOFTWARE MODULES IMPLEMENTING FW, IDS, NAT AND PROXY.

	Maximum Processing Capacity ($C^{pr,max}$)	Number of cores allocated (n^c)
FW	0.9 Gb/s	4
IDS	0.6 Gb/s	8
NAT	0.9 Gb/s	2
Proxy	0.9 Gb/s	4

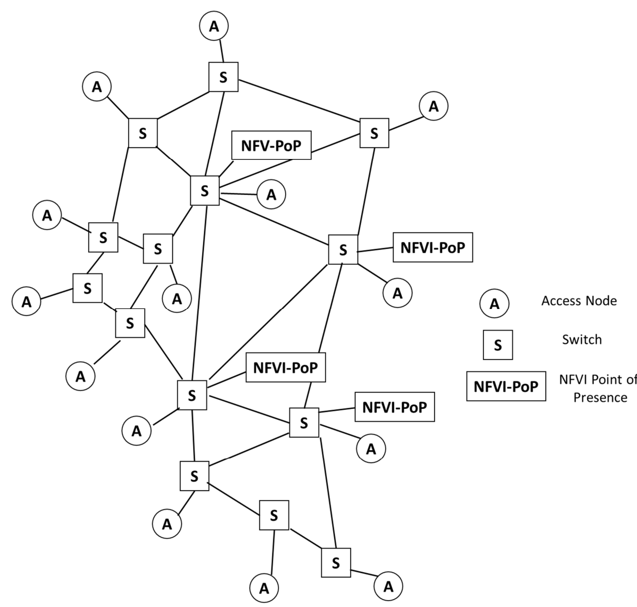

Figure 4. Deutsche Telekom network with four NFVI-PoPs.

We illustrate in Fig. 6 the SRCBA execution time as a function of the number of SFC requests for the Traditional Orchestration (TO) and Scalable Orchestration (SO) cases. We

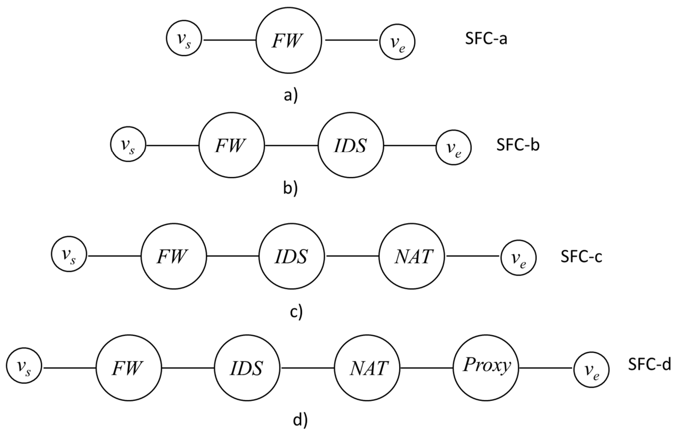

Figure 5. Composition of the Service Function Chain; four types are considered: the first one composed by a Firewall (FW) (a), the second one composed by a FW and Intrusion Detection System (IDS) (b); the third one composed by a FW, an IDS and a Network Address Translator (NAT) (c); the fourth one composed by a FW, an IDS, a NAT and a Proxy (d).

report the results when SFCs with one SF (SFC-a), two SFs (SFC-b), three SFs (SFC-c) and four SFs (SFC-d) of Fig. 5 are offered. We can observe how our proposed scalable orchestration allows for a remarkable reduction in SCRBA execution time. As a matter of example, when 2500 SFCs with FW and IDS are offered, SCRBA execution time of 24195 ms and 2475 ms are achieved for TO and SO respectively.

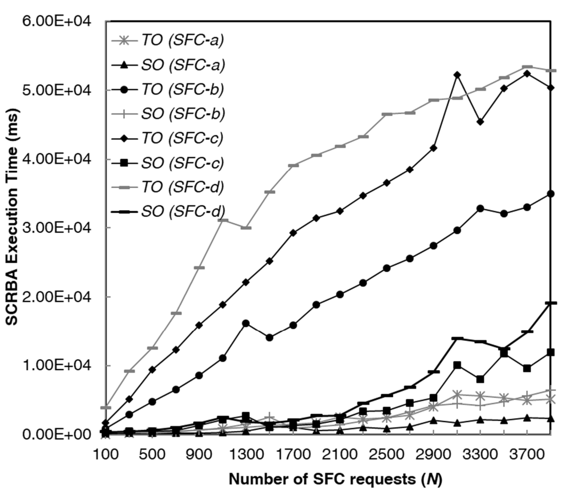

Figure 6. SCRBA execution time versus the number of SFC requests for the TO and SO solutions.

IV. CONCLUSIONS

A scalable resource orchestration solution for Multi-Provider NFV environments has been proposed and investigated. An SR-based data/control plane for the proposed solution has also been introduced. We have shown how the proposed solution allows for a drastic reduction in execution time of the SFC Routing and SFC Routing and Cloud and Bandwidth resource Allocation algorithm executed by an NFV orchestrator. In the analyzed study case we have verified a reduction by 90% of the execution time with respect to a traditional orchestration solution.

REFERENCES

[1] V. Eramo, A. Tosti, E. Miucci, "Server Resource Dimensioning and Routing of Service Function Chain in NFV Network Architectures", *Journal of Electrical and Computer Engineering*, Volume 2016 (2016), Article ID 7139852.

[2] V. Eramo, M. Ammar, F. G. Lavacca, "Migration Energy Aware Reconfigurations of Virtual Network Function Instances in NFV Architectures, *IEEE Access*, vol. 5, pp. 4927 - 4938.

[3] Binayak Kar, Eric Hsiao-Kuang Wu, Ying-Dar Lin "Energy Cost Optimization in Dynamic Placement of Virtualized Network Function Chains," *IEEE Transactions on Network and Service Management,* vol. 15, March 2018, pp. 372-386.

[4] F. Carpio, S. Dhahri, A. Jukan, "VNF Placement with Replication for Load Balancing in NFV Networks," in *Communications (ICC), 2017 IEEE International Conference on*, May 2017, Paris (France).

[5] Huan Chen, Xiong Wang, Yangming Zhao, Tongyu Song, Yang Wang, Shizhong Xu, Lemin Li, "MOSC: a method to assign the outsourcing of service function chain across multiple clouds, *Computer Networks*, vol. 133, March 2018, pp. 166-182.

[6] Chih-Chiang Wang, Ying-Dar Lin, Jang-Jiin Wu, Po-Ching Lin, Ren-Hung Hwang, "Towards Optimal Resource Allocation of Virtualized Network Functions for Hierarchical Datacenters, *IEEE Transactions on Network and Service Management*, available at https://ieeexplore.ieee.org/document/8424500/.

[7] B. Yi, X. Wang, K. Li, S.K. Das, M. Huang "A comprehensive survey of Network Function Virtualization, *Computer Networks*, vol. 133, March 2018, pp. 212-262.

[8] M. F. Bari, S. R. Chowdhury, R. Ahmed, R. Boutaba, O. C. Muniz, B. Duarteyh, "Orchestrating Virtualized Network Functions," *IEEE Transactions on Network and Service Management,* vol. 13, December 2016, pp. 725-739.

[9] N. F. Saraiva de Sousa, D. A. Lachos Perez, R. V. Rosa, M. A. S. Santos, C. E. Rothenberg, "Network Service Orchestration: A Survey," *Cornell University Library*, available at https://arxiv.org/abs/1803.06596

[10] Mijumbi R., Serrat J., Gorricho J., Bouten N., De Turck F., Boutaba R., "Network Function Virtualization: State-of-the-art and Research Challenges, *IEEE Communications Surveys and Tutorials*, vol. 18, First Quarter 2016, pp. 236-262.

[11] Y. Wang, P. Lu, W. Lu, Z. Zhu, "Cost-Efficient Virtual Network Function Graph (vNFG) Provisioning in Multidomain Elastic Optical Networks," *Journal of Lightwave Technology,* vol. 35, July 2017, pp. 2712-2723.

[12] A. Cianfrani, M. Listanti, M. Polverini, "Incremental Deployment of Segment Routing Into an ISP Network: a Traffic Engineering Perspective , *IEEE/ACM Transactions on Networking*, vol. 25, October 2017, pp. 3146 - 3160.

[13] F. Paolucci, "Network Service Chaining Using Segment Routing in Multi-Layer Networks, *Journal of Optical Communications and Networking*, vol. 10, June 2018, pp. 582 - 592.

[14] C. Filsfils, N. K. Nainar, C. Pignataro, J. C. Cardona, P. Francois, "The Segment Routing Architecture," in *2015 IEEE Global Communications Conference (GLOBECOM)*, December 2015, San Diego, CA, (USA).

[15] ETSI, "Network Functions Virtualisation Introductory White Paper, *SDN and OpenFlow World Congress*, https://portal.etsi.org/nfv/nfv white paper.pdf, Oct. 2012.

[16] V. Eramo, E. Miucci, M. Ammar, "Study of Reconfiguration Cost and Energy Aware VNE Policies in Cycle-Stationary Traffic Scenarios", *IEEE Journal Selected Areas of Communications*, vol. 34, April 2016, pp. 1-17.

[17] V. Eramo, E. Miucci, M. Ammar, "Study of Migration Policies in Energy-Aware Virtual Router Networks", *IEEE Communications Letters*, vol. 18, November 2014, pp. 1919-1922.

[18] Fei X., Liu F., Xu H., Jin H., "Adaptive VNF Scaling and Flow Routing with Proactive Demand Prediction, in *INFOCOM 2018 - IEEE Conference on Computer Communications, IEEE*, April 2018, Honolulu, HI (USA).

[19] X. Cheng, Y. Wu, G. Min, A. Y. Zomaya, "Network Function Virtualization in Dynamic Networks: A Stochastic Perspective", *IEEE Journal Selected Areas of Communications*, available at https://ieeexplore.ieee.org/document/8463557

[20] Po-Ching Lin, Ying-Dar Lin, Cheng-Ying Wu, Yuan-Cheng Lai, Yi-Chih Kao, “Balanced Service Chaining in Software-Defined Networks with Network Function Virtualization, *Computer*, vol. 49, November 2016, pp. 68 - 76.

[21] V. Eramo, F. G. Lavacca, “Computing and Bandwidth Resource Allocation in Multi-Provider NFV Environment, *IEEE Communication Letters*, vol. 22, October 2018, pp. 2060-2063.