MD2-NFV: The Case for Multi-Domain Distributed Network Functions Virtualization

Raphael Vicente Rosa, Mateus Augusto Silva Santos, Christian Esteve Rothenberg
Information & Networking Technologies Research & Innovation Group (INTRIG)
University of Campinas (UNICAMP), Sao Paulo, Brazil
Email: {raphaelvrosa,msantos,chesteve}@dca.fee.unicamp.br

***Abstract*—Motivated by the vision of Network Functions Virtualisation (NFV) spanning different administrative domains, this position paper makes the case for multi-domain, distributed NFV (MD2-NFV). To this end, we present MD2-NFV as a natural evolution of the NFV paradigm to deliver a distributed deployment of Virtualized Network Functions (VNFs) as a service. By means of three motivating use case scenarios, we discuss potential benefits and identify challenging features towards enabling advanced peering relationships between NFV domains.**

I. Introduction

Network Functions Virtualization (NFV) [1] has entered the networking scene promising to revolutionize how networks are built and operated by implementing networking devices as virtualized software-only appliances using commodity hardware. As performance and scalability challenges to realize the NFV vision are being sorted out, a number of issues remain open regarding the management and orchestration realms of NFV infrastructures [2].

One of the tenets of NFV is to consider end-to-end network services as a set of connected or chained Network Functions (NFs) delivered by virtual and/or physical appliances. Examples of NFs include firewalls and load balancers, commonly delivered via middleboxes deployed in today's networks, but also more basic functional blocks such as filtering, routing/switching, traffic shaping, and so on. Instead of having these NFs as static entities processing all traffic at strategic network choke points, NFV promises the means to allow a fine-granular, dynamic composition of these NFs in the form of Network Service Chains (NSCs) [3] that can be represented as Virtual Network Functions Forwarding Graph (VNF-FG). There is growing interest of the research community in simplifying the management of NFs (e.g., [4]–[6]) by turning their connectivity programmable, flexible (reconfigurable to meet demands) [7], and efficient (avoiding path inflation) [8].

Interest in changing the state of affairs is shared by telecommunication operators in an organized effort within ETSI NFV Industry Specification Group (ISG) to convey new ways to design, deploy, and manage networking services. There is a natural synergy between Software Defined Networking (SDN) [9] and NFV –both advocate the use of software components in commodity hardware and share most of the strategic objectives (e.g., innovation, reduced OPEX & CAPEX, new business models). One clear distinction is the focus of SDN in separating the control and data planes and deliver appropriate programming abstractions. As such, SDN and NFV are regarded as mutually beneficial. More specifically, SDN is being proposed as a NSC facilitator by providing adequate mechanisms to steer traffic flows through a coordinated set of virtual NFs (VNFs) (e.g., [10], [11]).

978-1-4799-5804-7/15/$31.00

NFV naturally introduces the case for decoupling network functions from location. In turn, similarly to the cloud computing model, untangling VNFs from the actual resource pool (server, storage, networking) where they are executed allows carriers to build an adequate NFV Infrastructure (NFVI) to deliver VNF as a service (VNFaaS) [12]. At this point, the optimization of VNF location within the different end-to-end location options (e.g., customer premises, service/aggregation PoP, carrier data centers) becomes appealing for multiple reasons [13], including maximizing the Quality of Experience (QoE) by bringing VNFs closer to users, or consolidating more VNFs to reduce costs from the underlying infrastructure. The resulting distributed model where VNFs are instantiated at the customer premises has been referred to as Distributed NFV (D-NFV) [14].

In this research-oriented position paper, we explore a natural step ahead by considering the case of multiple administrative domains (MD2-NFV: Multi-Domain Distributed Network Functions Virtualization). Along that journey, we identify a series of functional goals, including the interaction with SDN as an enabling technology for VNFs interconnection across edge networks, resulting in a series of open research questions. We first introduce the rationale behind MD2-NFV and its benefits (Section 2). We then present three motivating use cases (Section 3) that suggest MD2-NFV can deliver innovative services with high-performance and reliability through the collaboration of different administrative domains peering together beyond traditional customer-provider relationships. Finally, we put together the identified challenges (Section 4) that correspond to some of own research perspectives around the realization of MD2-NFV.

II. MD2-NFV

As shown in Figure 1, one of the main characteristics of the MD2-NFV case is to allow the hosting of VNFs in arbitrary locations (e.g., data centers, aggregation nodes, customer premises) from multiple NFVI providers (administrative domains). The case follows the rationale that a VNF should be placed at the network location where it can best meet its multi-dimensional operational requirements (e.g., processing power, network I/O and end-to-end performance). For instance, firewalls, IDS/IPS, and NATs are best placed at (or as close as possible to) Customer Premises Equipments (CPEs). Being

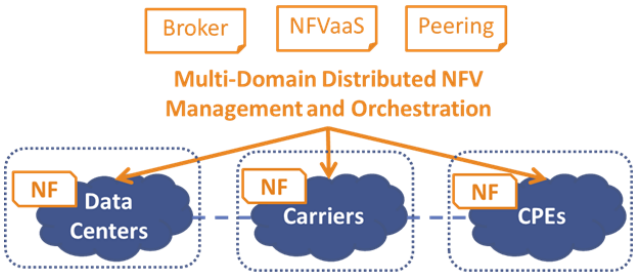

Fig. 1: Overview of multi-domain distributed NFV.

able to select the VNF location from multiple candidate Network Points of Presence (N-PoPs) allows not only optimizing in terms of processing and network performance (low latency, high bandwidth) but also guaranteeing high resilience by replicating VNFs that avoid shared risk factors.

The required overall manageability and flexibility in the placement of VNFs is expected to be supported by the so-called NFV Management and Orchestration (NFV-MANO) [15], a functional entity in the ETSI NFV ISG architectural work that provides VNFs with the required resources such as computing, networking and storage. The orchestration framework shall allow VNFs to be placed at the desired NFVI locations and have their interconnection path defined through adequate NSC mechanisms [3].

The adequacy of MD2 NFVIs would also allow the migration of applications and services from centralized to distributed outlooks. Besides, load-balancing across VNF instances could create high availability of services scaling in proportion to the underlying resource pool. Other advantages include VNFs snapshotting and commissioning performance test/birth certificates to reduce OPEX (e.g., extraction of metrics for proportional offering and charging of VNFs resources). Altogether, MD2-NFV seeks to offer the following benefits.

Performance: placing network functions near customers contribute to higher performance, QoS, and scalability targets.

Agility: NFs and NSCs deployment, monitoring, and maintenance tasks can be automated with multi-domain interactions of MANO entities, performing dynamic and flexible automatic/on-demand customer services provisioning.

Resilience: NFs can be mirrored in different NFVIs as a redundancy service assurance method, including migration to maintain transparent service continuity.

Policy: network service driven applications can be deployed on customized NFVI-PoPs and negotiated with customers to improve SLA through agile policies.

Cost: through economies of scale and dynamics of NFV marketplaces with multiple NFVI providers, CAPEX and OPEX reduction is expected (similarly to the cloud) by pursuing a shared model of multi-domain, distributed NFVI-PoPs.

Faster service delivery by chaining VNFs across administrative boundaries adds to the NFV list of challenges the ability to negotiate NSCs while satisfying multiple SLA requirements (e.g., placement, performance, availability). SDN appears as an enabling technology to abstract details from each domain and provides the means to programmatically set the required state in the data plane to realize the negotiated connectivity among the VNFs, probably combining actuation on both physical and virtual data plane elements. However, in addition to the intrinsic challenges of SDN due to scalability issues and technology maturity concerns, adequate horizontal interfaces (often referred to as east-westbound APIs) between SDN controllers need still to be defined.

One of the multiple open design questions to realize MD2-NFV is the question whether (*i*) leveraging horizontal interfaces between SDN controllers to support the envisioned MD2-NFV use cases, or (*ii*) moving the inter-domain communication problem up to the role of an overarching orchestrator acting as a broker, which in turn, will use services (northbound APIs) from SDN controllers. Initial related work on multi-domain-layered service orchestration [16] suggests the latter option. However, the degree of common northbound APIs expected from SDN controllers at different domains remains unclear.

Our vision on the future of SDN-enabled MD2-NFV services is at the cross roads of the evolution of Software-Defined eXchanges (SDX) [17]. The datacenter-like infrastructure of Internet eXchange Points (IXP) and their distributed, multi-domain nature turns them as candidate NFVIs enabling a rich marketplace and ecosystem around MD2-NFV use cases.

III. Motivating Use Cases

This section outlines three use case scenarios that motivate, through value-added examples, the case for MD2-NFV. The first use case presents how NFVI providers may negotiate virtualized resources for VNFs to be orchestrated and interconnected defining NSCs across different domains. The second use case shows multi-level, inter-domains bandwidth allocation through peering negotiations to fulfill NSC requirements. Finally, the third use case highlights service continuity views for NSCs and VNFs availability metrics to supply fault tolerance assurance in inter-domain NSCs.

In the proposed use case methodology, a context and goals are first presented before describing a set of actions expected from the MANO layer, including relevant events and information processing. Finally, outputs are described with regard to success conditions and failed end protections.

A. MD2-VNFs Management and Orchestration

The approach of having VNFs close to the customer allows optimized geographical distribution and dynamic deployment of NSCs. NSCs can be deployed using different resources (e.g., compute, storage, network), virtualization levels (e.g., container, bare metal), and NFVI hosting environments (e.g., data centers, CPEs), in such a way that fast network services provisioning shall be possible at different granularities.

The business model behind this use case stands for distributed NSCs across multiple domains to provide low latency, high resilience, and flexible customer services in distributed N-PoPs. Even though not specified here, peering mechanisms are required to exchange information between the different actors (NFVI providers and consumers) following the model of NFVaaS, i.e. jointly providing IaaS and NaaS tailored for NFV. The result of such kind of negotiation should include resource

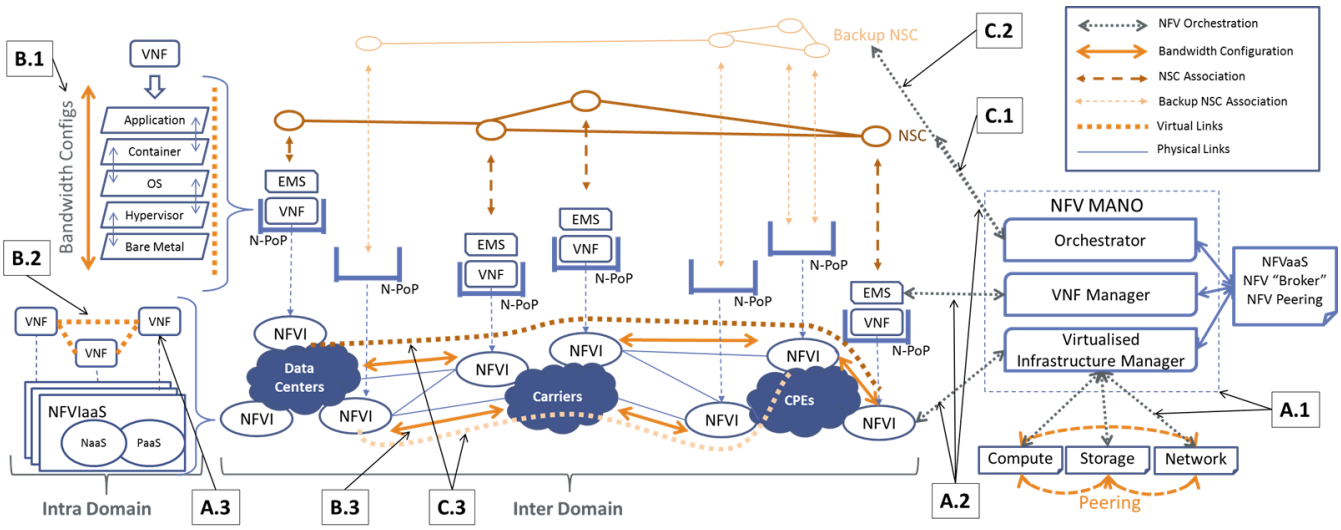

Fig. 2: Multi-Domain Distributed NFV Use Case Scenarios.

allocation for NSCs (e.g. VNFs placement) and flow steering rules to interconnect those VNFs. To achieve this goal, heterogeneous domains need to exchange information regarding their NFVI capabilities and the requested VNFs. Henceforth, extending the models of resource negotiation in cloud computing (client-provider), MD2-NFV peering (provider-provider) are called for through the management and orchestration plane of each interested party, which could include specific and general domain policies.

Goals: (*i*) Realize NSCs based on distributed VNFs, structured and established in N-PoPs across multi-domain environments resource negotiation; (*ii*) Negotiation of NSCs as well as the actual placement of VNFs takes place in the form of weather maps (e.g., an infrastructure map of where the allocation of VNFs are feasible based on the type of resources available) according to a joint provisioning of NFVaaS in multi-domain environments; (*iii*) Service Level Agreements dictate the relationships of NSCs accomplished with on-demand and/or automatic mechanisms between management and orchestration planes of each domain.

Actions: **(A.1)** process inter-domain NSCs establishment/requisition, update domain resources, and negotiate VNFs parameters and placement information; **(A.2)** perform multi-objective VNFs placement and interconnection; **(A.3)** VNFs weather maps and placement suggestions (including control plane VNFs) (see Figure 2).

Outputs: scheduled VNFs placement for on-demand/automatic provisioning of NSCs; (set of) paths to establish and configure NSCs requirements; and updates of infrastructure domain state based on NSC resources allocation.

Success conditions: when established, a fully configured NSC results in (*i*) VNFs deployed across different domains, and (*ii*) updated resources in the participating infrastructures based on-demand and/or automatic provisioning systems. **Failed end protection:** an on-demand and/or automatic cleaning procedure of VNFs state, saving the domain state and configuration changes when/where required.

B. Distributed Bandwidth Negotiation

Bandwidth negotiation across different network domains has been a long standing challenge for network operators, with different views of bandwidth contracts, where many carriers use a diverse set of techniques according to multiple parameters such as geo-location, percentiles of usage, sampling mechanisms, and so on. When MD2-NFV enters the scene, a number of requirements arise for its broad realization including VNF portability and guaranteed performance proportional to the underlying available resources. The portability of VNFs can demand large variations of bandwidth usage at different network stack levels (e.g., OS, NIC, hypervisor). The proper identification and classification of VNFs bandwidth resources usage, including initial and running configurations, induces the possibility of VNFs bandwidth negotiation scenarios (in spirit of ISP peering) to accommodate mobile VNFs and NSCs across different network domains.

Provided the ability of deploying probe entities capable of extracting information of N-PoPs regarding their current state of resources at distinct network and processing levels from different domains, this use case highlights network performance negotiation for different VNF requirements and granularities prior to the allocation and/or sharing of bandwidth necessary to interconnect D-NFV instances.

The actors of this use case shall be able to define strict bandwidth usage limitations by means of interfaces between SDN controllers and the NFV-MANO plane. Based on the existence of such negotiations, dynamic, granular bandwidth contracts at different network and virtualization stack layers (e.g., OS, hypervisor) and environments (e.g., intra- and inter-domain) shall be possible, overcoming limitations of existing bandwidth usage metering and auditing techniques.

Goals: (*i*) perform distributed bandwidth negotiation for the VNFs in single and multiple domains. The structure of VNFs deployment, the running configuration (in terms of bandwidth requirements), and its environment will compose the main factors for local and distributed VNF bandwidth negotiation; (*ii*) achieve proper organization of the VNFs

communication links (e.g., NICs, bridges, virtual switches); (*iii*) extract information of usage according to on-demand and automatic resource configuration for bandwidth management and consistent auditing purposes.

Actions: (B.1) negotiate of VNFs bandwidth requirements in different levels; **(B.2)** compute and store current infrastructure bandwidth resources (multiple domains, environments and levels); **(B.3)** manage links interconnecting VNFs to meet bandwidth requirements (see Figure 2).

Outputs: bandwidth defined links at different network levels (e.g., NetMap, NIC); QoS flow entries using metering and/or queue actions (e.g., OpenFlow meter tables) based on the target link characteristics negotiated for the distributed VNFs; updates of current infrastructure state based on VNF-FGs resource allocation.

Success conditions: the negotiation of bandwidth incurs on the configuration of network links and installation of flow rules across different network domains and environments as negotiated. **Failed end protection:** a response to the management entity containing alternative link instance possibilities with the respective bandwidth requirements demanded or optional links with alternative resources.

C. NSCs Reliability

Similar to cloud computing, there are business concerns when sharing resources among different NFVI providers. An underlying challenge is the ability to specify fault tolerance network parameters in the NFV service provider contract. In a dynamic environment –with many VNF-FGs and where mobile VNFs could be ported to different VNFI providers– the specification of NSCs service continuity levels and resilience negotiations across different domains is required.

This use case intends to seek means for the specification of VNFs and VNF-FGs availability requirements as well as enable their negotiation across multiple network domains and environments. It also aims at deploying NSCs with zero or minimal measurable impact regarding their service assurance as well as providing suitable parameters to define service continuity metrics for NSCs and VNFs. Such goals can be supported by idle infrastructure resources available to maintain NSCs and their corresponding VNFs operational according to availability metrics. This requires the ability to negotiate and allocate idle fault tolerance resources from single or multiple domains to be used when required. For example, when a VNF of a critical NSC fails, it needs to be transparently and automatically instantiated to guarantee the service continuity of the NSC. In this case, the NFVI provider can negotiate spare resources to attend zero/measurable impact, automatic or on-demand fault tolerance for different guarantee levels (e.g., 1+1, N+1).

By developing adequate measurement mechanisms and defining adequate metrics, it shall be possible to establish service continuity in different levels of operation for entities such as NSCs, VNFs, network infrastructures, hypervisors, etc. Evolving this idea, in the long term, with extracted data from failure events of NSCs and VNFs and the correlation of all the metrics associated with service assurance of NSCs, it may be possible to infer metrics for full end-to-end NSCs service continuity and VNFs resiliency. Leveraging this knowledge, service and infrastructure providers can offer and negotiate network services assurance/continuity in different granularities, levels, environments, and pricing models.

Goals: (*i*) define a clear VNFs resilience metrics related to NSCs service continuity; (*ii*) calculate service continuity parameters for NSCs and their associated VNFs based on availability requirements; (*iii*) create mechanisms to negotiate common fault tolerance requirements between multiple-domains based on NSCs and VNFs service continuity metrics.

Actions: (C.1) negotiation and classification of VNFs reliability based on performance parameters (e.g., packet loss, latency); **(C.2)** compute current and backup infrastructure provider fault tolerance resources; **(C.3)** perform multi-objective VNFs resilience computation and optimized deployment in available infrastructures (see Fig. 2).

Outputs: requisitions of VNFs virtual instrumentation functions and metrics (e.g., packet timestamps); updates of current infrastructure state based on VNF reliability and resources allocation; infrastructure and NSCs service continuity classifications based on minimal impact.

Success conditions: the negotiation of MD2-NFV fault tolerance parameters incurring in well-provisioned service continuity indexes (e.g., VNFs availability metrics) with zero/measurable impact negotiated between multiple network domains and environments. And based on the requirements, the storage of the network state and configuration (e.g., required for auditing or reliability analytics). **Failed end protection:** a reestablishment of VNFs availability classification and renegotiation of service continuity parameters.

IV. Challenges and Research Directions

Upon introducing a selection of MD2-NFV use cases, we now present a set of challenges when attempting to spread VNFs through different network locations in a distributed deployment of VNFs as a service.

Edge infrastructure evolution: new kinds of virtualization layers are currently being introduced (e.g., Netmap [18], OpenDataPlane), which demand hardware and software features to run at the same time yielding new performance capabilities. Besides, the place for VNFs management system still needs its requirements to coexist as a VNF management and orchestration interface. As exist today, end users equipments (e.g., DSL modems, setup boxes) may be replaced by commodity hardware with dumb thin clients enabling orchestration interfaces with Service Provider network operators.

Connected-centered NFV-MANO: heterogeneous resource (e.g., computing, storage, network) peering-like negotiation is a central feature to fulfill MD2-VNFs requirements. Different forms to connect network, storage and computing realms are necessary, including well-defined interfaces, since VNFs scaling tasks shall be capable of scheduling virtualized resources on different levels and environments (e.g., hypervisor, OS). The underlying hardware resources need to be managed by a connected-centered NFV-MANO, a especially critical issue in multi-domain scenarios.

NFV-MANO interfaces to SDN: in contrast with the centralized nature of NFV inside a single datacenter, the flow

coordination among distributed VNFs requires horizontal interfaces between administrative domains which may consider the capabilities of (heterogeneous) SDN controllers. A related challenge is lacking of well-defined interfaces between components that realize VNFs deployment and their interconnection.

Placement of control plane: in the case of SDN-enabled networks, logically distribution of control plane is still under on-going research, making SDN today mostly applied to intra-domain scenarios such as data centers. Service Providers dispose micro data center at their network borders, including IXPs (e.g. coped with SDN [17]) that could be used as NFVI-PoPs for both data and control plane VNFs. These locations could also be regarded as proxies for centralized control office in different hierarchical arrangements (e.g., master-slave, nested).

Resources capabilities discovery: with VNFs executing on shared environments, the qualification of resources available on NFVI-PoPs should be measurable regarding to computing, storage and network domains in automatic/on-demand ways. Discovering MD2-NFV infrastructure idle resources incurs on evolved cloud computing methods, e.g., scheduling of NSCs requires bandwidth and latency metrics collected through distributed measurement mechanisms.

Monitoring and auditing mechanisms and metrics: in the sense to provide service assurance levels to customers as it exists nowadays, NFVI providers should implement the means to monitor distributed VNFs which could demand scalability challenges based on network overhead, NFV-MANO controller and VNFs performance and NFVI capabilities. In addition, availability metrics necessary to those tasks should be created and fault tolerance mechanisms should be implemented to deliver confidence level guarantees.

Policies: representational policy languages for specific application domains (e.g. OpenStack congress) are still to be investigated for MD2-NFV. It should be possible to express storage, compute, and network policies per domain. However, the orchestration and management of distributed VNFs need some level of generic and independent domain policies to allow heterogeneous implementations inter-working for wide auditing, monitoring, and policy enforcement purposes.

Security: Access control to the main components of a NFVI opens security vulnerabilities that become more critical across domain boundaries. Role based access purposes could allow customers customization of VNFs and NFVI-PoPs according to their privileges and needs while restricting intruders by securing NFV-MANO communication channels.

Business model: with the introduction of network services orchestration in MD2-NFV, VNFs may be moved around different environments being held by multiple providers on their own or rented NFVIs. Entire NSCs themselves may recursively follow the same operational shift. Incentives for NFVI providers to share their infrastructure should be provided, as well as for new pure software companies, which may evolve from developers of VNFs and platforms for NFVI-PoPs to aggregators of NFVaaS providers and NSCs brokers, following business dynamics similar to service models in the cloud.

V. Conclusion

This paper states for the MD2-NFV case. We introduced this concept as a natural evolution of the NFV broadening use cases combined with lessons learned from the cloud computing industry. We presented three exemplifying use cases that motivate a series of features for the realization of MD2-NFV. The identified missing pieces in NFV and SDN developments are only part of the challenges and open research directions, emerging from the multi-domain extensions of NFV, that distributed environments are likely to face.

VI. Acknowledgments

This work was supported by the Innovation Center, Ericsson Telecomunicaes S.A., Brazil.

References

[1] NFV White Paper, "Network Functions Virtualisation: An Introduction, Benefits, Enablers, Challenges & Call for Action. Issue 1," Oct 2012.

[2] ETSI, "Network functions virtualisation requirements," ETSI ISG, Tech. Rep. GS NFV 004 v1.1.1, October 2013.

[3] W. John et al., "Research directions in network service chaining," in *IEEE SDN4FNS*, 2013, pp. 1–7.

[4] A. Gember, P. Prabhu, Z. Ghadiyali, and A. Akella, "Toward software-defined middlebox networking," in *Proceedings of the 11th ACM HotNets*, 2012, pp. 7–12.

[5] Z. Qazi, C.-C. Tu, R. Miao, L. Chiang, V. Sekar, and M. Yu, "Practical and incremental convergence between sdn and middleboxes," *Open Network Summit, Santa Clara, CA*, 2013.

[6] Y. Zhang et al., "Steering: A software-defined networking for inline service chaining," in *Proceedings of the 21st IEEE ICNP*, 2013, pp. 1–10.

[7] A. Gember-Jacobson, R. Viswanathan, C. Prakash, R. Grandl, J. Khalid, S. Das, and A. Akella, "OpenNF: Enabling innovation in network function control," in *Proceedings of the 2014 ACM Conference on SIGCOMM*, 2014, pp. 163–174.

[8] B. Anwer, T. Benson, N. Feamster, D. Levin, and J. Rexford, "A slick control plane for network middleboxes," in *ACM SIGCOMM HotSDN*, 2013, pp. 147–148.

[9] D. Kreutz, F. M. V. Ramos, P. Veríssimo, C. E. Rothenberg, S. Azodolmolky, and S. Uhlig, "Software-Defined Networking: A Comprehensive Survey," *Proceedings of the IEEE*, vol. 103, no. 1, p. 63, 2015. [Online]. Available: http://arxiv.org/abs/1406.0440

[10] S. K. Fayazbakhsh, V. Sekar, M. Yu, and J. C. Mogul, "Flowtags: Enforcing network-wide policies in the presence of dynamic middlebox actions," in *ACM SIGCOMM HotSDN*, 2013, pp. 19–24.

[11] J. Blendin, J. Rückert, N. Leymann, G. Schyguda, D. Hausheer, "Software-Defined Network Service Chaining." in *Third EWSDN*, http://ewsdn.eu, 2014.

[12] T-NOVA, "Network Functions as-a-Service over Virtualised Infrastructures," http://cordis.europa.eu/fp7/ict/future-networks/documents/call11projects/t-nova.pdf.

[13] P. Sköldström et al., "Towards unified programmability of cloud and carrier infrastructure," in *Third EWSDN*, http://ewsdn.eu, 2014.

[14] ETSI, "Network functions virtualisation use cases," ETSI ISG, Tech. Rep. GS NFV 004 v1.1.1, October 2013.

[15] ——, "Network functions virtualisation architectural framework," ETSI ISG, Tech. Rep. GS NFV 002 v1.1.1, October 2013.

[16] A. Csomaet al., "Multi-layered service orchestration in a multi-domain network environment. Demonstation." in *Third EWSDN*, http://ewsdn.eu, 2014.

[17] A. Gupta et al., "Sdx: A software defined internet exchange," in *Proceedings of ACM SIGCOMM*, 2014, pp. 579–580.

[18] L. Rizzo, G. Lettieri, and V. Maffione, "Speeding up packet i/o in virtual machines," in *Proceedings of ACM/IEEE ANCS*, 2013, pp. 47–58.