Availability Evaluation of Service Function Chains Under Different Protection Schemes

Jean-Pierre H. Asdikian[1], Leila Askari[2], Omran Ayoub[3], Francesco Musumeci[2], Stefano Bregni[2], Massimo Tornatore[2]
[1]*American University of Science and Technology, Beirut, Lebanon.* [2]*Politecnico Di Milano, Milan, Italy.* [3]*SUPSI, Lugano, Switzerland*

***Abstract*—Network Function Virtualization (NFV) calls for a new resource management approach where virtualized network functions (VNFs) replace traditional network hardware appliances. Thanks to NFV, operators are given a much greater flexibility, as these VNFs can be deployed as virtual nodes and chained together to form Service Function Chains (SFCs). An SFC represents a set of dedicated virtualized resources deployed to provide a certain service to the consumer. One of its most important performance requirements is availability. In this paper, the availability achieved by SFCs is evaluated analytically, by modelling several protection schemes and given different availability values for the network components. The cost of each protection scheme, based on its network resource consumption, is also taken into account. Extensive numerical results are reported, considering various SFC characteristics, such as availability requirements, number of NFV nodes and availability values of network components. The lowest-cost protection strategy, in terms of number of occupied network components, which meets availability requirement, is identified. Our analysis demonstrates that, in most cases, resource-greedy protection schemes, such as end-to-end protection, can be replaced by less aggressive schemes, even when availability requirements are in the order of five or six nines, depending on the number of elements in the service function chain.**

***Index Terms*—Service Chaining, Virtual Network Functions, Protection, Availability.**

I. Introduction

To support new emerging 5G services, such as remote diagnosis and smart factory, network operators are expected to guarantee unprecedented availability requirements. Thanks to recent advances in Network Function Virtualization (NFV), these services are realized by concatenating software instances, called Virtual Network Functions (VNFs), forming a Service Function Chain (SFC). A SFC is considered available only if all of its computational (VNFs) and transmission (communication links) components are available.

Various protection techniques can be employed to guarantee availability of SFC components in case of failure. In particular, an *end-to-end* (E2E) protection strategy, i.e., protecting all components along a SFC, is adopted to guarantee highest availability levels. However, for some SFCs, employing end-to-end protection might be more than required (i.e., availability requirements can be met with less aggressive protection techniques), resulting in resource over-provisioning.

Therefore, is it essential to employ a protection strategy that replicates just enough SFC components to meet availability requirements, while averting exaggerated occupation of network resources and unnecessary Operational Expenditure (OpEx).

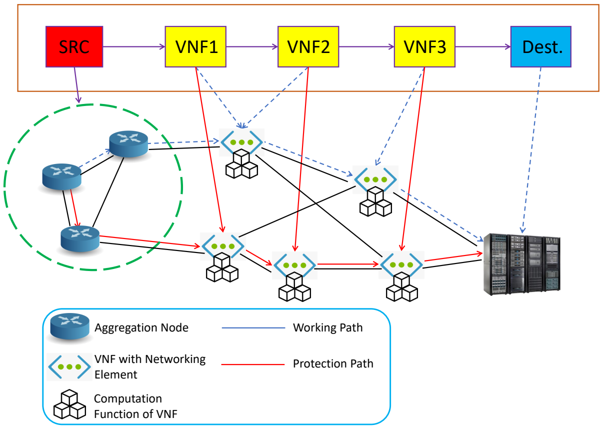


Fig. 1. Example of Network Model

Figure 1 shows an example of protecting a SFC made of 3 VNFs, by provisioning a primary working chain (solid red line) and a backup chain (dashed blue lin). An SFC is represented by a set of VNFs (virtual nodes) and a set of links between them (virtual links). The VNFs are mapped onto physical nodes equipped with computational resources and therefore capable of hosting VNFs, while virtual links between VNFs are mapped onto physical links. In this example, the SFC is protected by an end-to-end protection strategy, where all SFC components are replicated with the working and the backup chains being completely disjoint. Therefore, the overall amount of network resources allocated to provide end-to-end protection is 6 VNFs and 9 links.

Instead of replicating all SFC components, other protection strategies have been studied, which protect a subset of the SFC components, for instance virtual link protection and virtual node protection, which are considered less aggressive than end-to-end protection. As it will be detailed in Sec. III, these other protection strategies use network resources differently, hence providing different levels of protection.

It is worth remarking that, as it will be demonstrated later in this paper, not necessarily the more aggressive the protection scheme is, the higher is the protection level provided, as it strictly depends on the availability of network resources. In this context, it is not trivial to determine what protection scheme should be considered for a SFC, i.e., which SFC components should be protected to meet the availability requirement whilst limiting network resource occupation. In fact, the optimal choice depends on the number of service functions of the SFC, its availability requirement and the availability of network components.

Putting it simply, the question to answer is the following: given a service chain, availability of switching nodes, availability of links and availability of computing nodes, *to meet some given availability requirement for an SFC, what protection strategy should be employed?*

In this work, we address this question by providing closed-form analytical expressions of the availability achieved by different protection schemes for a given SFC. We assign a cost to each protection scheme based on network components needed to guarantee protection, and consider as the optimal protection scheme the one that meets availability requirements with minimum cost. Analytical results show that an end-to-end protection method can be substituted by a less-aggressive protection scheme that replicate most vulnerable components (components with relatively low availability) in the network to meet availability requirements of SFC.

The paper is organized as follows. In Sec. II we survey works focusing on protection and availability-aware provisioning strategies of SFC. In Sec. III we present the protection strategies considered in our study, their analytical model and the tool developed to calculate the availability of a SFC. Sec. IV reports and discusses analytical numerical results. Finally, Sec. V concludes the paper.

II. Related Work

Many works investigated the problem of SFC provisioning. Ref. [1] focuses on mapping links and nodes on a shared substrate network with the aim to allocate multiple virtual network requests considering node and link constraints. Ref. [2] dynamically allocates network resources that correspond to service chains instead of statically allocating them, to reduce both CapEx and OpEx without any QoS compromise. Ref. [3] targets dynamic Service Chaining, by showing space and time diversity in service chaining, with a higher degree of dynamism and flexibility with respect to conventional hardware based architectures. Ref. [4] focuses on dynamic SFC deployment by formulating an ILP on a practical network scenario while Ref. [5] proposes an ILP formulation to find the best feasible paths and virtual function placement.

Moreover, other works tackled the SFC provisioning problem considering availability requirements [6], [7] and protection [6], [8]–[11]. For instance, Ref. [7] proposes an approach to minimize cost of protection, however without considering the impact of link availability on SFC availability, while Ref. [6] proposes an availability-aware survivable virtual network embedding to satisfy the availability requirements of all virtual components in the network. Ref. [8] focuses on resilient SC provisioning by proposing different protection schemes while meeting SFC's latency requirements. Ref. [10] explores VNF placement strategies to satisfy SC availability requirements while minimizing bandwidth allocation and backup VNF utilization. Considering dynamic traffic, Ref. [16] proposes a genetic algorithms to provide end-to-end protection for a deployed SC by providing backups for all VNFs and links connecting them. Authors evaluate the performance of their algorithm in a Wavelength Division Multiplexing (WDM) network in terms of service blocking ratio comparing it with those of unprotected approach and the approach in which only VNFs of a SC are protected (VNF protection). In [17], authors devise a strategy to ensure availability of a SC using both physical network protection and virtual layer VNF replicas. By assessing how to distribute VNF replicas between the primary path and the backup path to achieve maximum availability of the SC, they decide the number of replicas in the SC for each VNF, and distribute the replicas to physical nodes while preserving sequentially among VNFs.

While many articles have tackled the availability and protection problem of SCs, to the best of our knowledge no work has inspected the problem from an analytical point of view, with the aim of realizing guidelines for choosing the protection scheme that meets availability requirements, based on availability of network components, while minimizing overall cost of components used.

III. Evaluation of Availability Achieved by Service Chain Protection

Availability is defined as the probability (equivalently, the expected fraction of time) that a service is available, i.e., it operates meeting given performance requirements. To offer a set of different solutions with different levels of availability, in this work we have considered the following protection strategies.

- Unprotected: no protection provided
- End-To-End (E2E) Protection: all components of the SFC are duplicated and located in failure disjoint locations
- Virtual-Link Protection: a protection strategy that provides backup failure-disjoint SFC links, but not backup SFC nodes
- Virtual-Node Protection: a protection strategy that provides backup failure-disjoint SFC nodes (i.e., VNFs are replicated in another node), but not backup SFC links
- Only Backup VNF: A backup VNF is created to the VNF in the same node
- E2E protection with Backup VNF: An end-to-end protection of links and virtual nodes with a backup for all VNFs
- Virtual-Link protection with Backup VNF: A backup VNF is created to the VNF along with protection to the links but not the nodes
- Virtual-Node protection with Backup VNF: In addition to virtual-node protection, a backup VNF is deployed for each VNF in the same node

In our analysis, we have assumed that a VNF placed at a virtual node is not shared among different SFCs. Moreover, in a single node, we have assumed that each computing and switching component has its own respective availability. Main abbreviations are listed in Table I.

TABLE I
LIST OF ABBREVIATIONS AND THEIR CORRESPONDING DEFINITION

Abbreviation	Definition
A_{SFC}	SFC availability
A_{PL}	Physical Link Availability
A_{VN}	Virtual Node Availability
A_{SW}	Switching Nodes Availability
W	Working (Path or Link)
B	Backup (Path or Link)
E	Event referring to a random variable depicting the probability of success/failure of an SFC
VNT	Virtual Network Terminal
WP	Protected Workpath
Meta	Self-defining node with only a switching task

A. Unprotected

The overall availability of an unprotected SFC A_{SFC_w} is simply given by the product of the availabilities of the SFC components, that is

$$A_{SFC_w} = \left[\left(\prod_{i=1}^{N} A_{VN_i} \right) \left(\prod_{j=1}^{M} A_{PL_j} \right) \left(\prod_{k=1}^{L} A_{SW_k} \right) \right] \quad (1)$$

B. End-to-End (E2E) Protection

In this scheme, all elements of the SFC are replicated along a backup path (neither nodes nor links are shared between the working and backup paths), providing resiliency against single-node and single-link failures. In Fig. 2, the left graph depicts an example, where 6 nodes and 9 links are utilized to provide an End-To-End protection for an SFC of 3 VNFs. The working path is highlighted in green and the backup path is highlighted in red. Then, the availability of a SFC with E2E Protection is given by:

$$P\{E_{SFC}\} = P\{E_{SFC_W} \cup (\overline{E_{SFC_W}} \cap E_{SFC_B})\} \quad (2)$$

where $P(E_{SFC_W})$ and $P(E_{SFC_B})$ represent the availability of the working path and the backup path, respectively. This is equivalent to:

$$A_{SFC} = A_{SFC_W} + (1 - A_{SFC_W}) A_{SFC_B} \quad (3)$$

The availability of the SFC working path is then given by

$$\begin{aligned} P\{E_{SFC_W}\} &= P\{E_{Link}\} \cap P\{E_{VN_1}\} \cap P\{E_{Link}\} \\ &\cap P\{E_{VN_2}\} \cap P\{E_{Link}\} \cap P\{E_{VN_3}\} \cap P\{E_{Link}\} \end{aligned} \quad (4)$$

where

$$P\{E_{VN_i}\} = P\{E_{VNF_i}\} \cap P\{E_{SW_i}\} \quad (5)$$

An analogous expression holds for the availability of the SFC backup path.

C. Virtual-Link Protection

This strategy provides protection solely to links, while nodes are not protected. Virtual nodes can be shared, but virtual links cannot (see Fig. 2, middle graph). Note that the VNFs have no protection in this case. The availability of SFC is then given by:

$$P\{E_{SFC}\} = P\{E_{vnt_1} \cap E_{vnt_2} \cap E_{vnt_3} ... \cap E_{vnt_f}\} \quad (6)$$

where each corresponding smaller network event is equal to:

$$P\{E_{vnt_1}\} = P\{E_{vnt_{1_W}} \cup (\overline{E_{vnt_{1_{WP}}}} \cap E_{vntp_{1_B}})\} \quad (7)$$

This is equivalent to

$$A_{vnt_1} = A_{vnt_{1_W}} + (1 - A_{vnt_{1_{WP}}}) A_{vntp_{1_B}} \quad (8)$$

D. Virtual Node Protection

This strategy focuses on backing up the virtual nodes of the SFC. Although the virtual links may be shared, the nodes on the main path and backup path are completely separate.

The calculation of overall availability here is based on the availability of *Meta* nodes:

$$P\{E_{SFC}\} = P\{E_{Meta_1} \cap E_{Meta_2} \cap E_{Meta_3} ... \cap E_{Dest_f}\} \quad (9)$$

$$A_{SFC} = A_{Meta_1} A_{Meta_2} A_{Meta_3} ... A_{Dest} \quad (10)$$

This means that the network is segmented in such a way, that a switching ("Meta") node is at the end of the sub-network (where each sub-network is linked to one another, creating the SFC network), and is then defined by:

$$P\{E_{Meta}\} = P\{E_{Meta_W} \cup (\overline{E_{Meta_{WP}}} \cap E_{Meta_p})\} \quad (11)$$

This is equivalent to

$$A_{Meta} = A_{MetaW} + (1 - A_{Meta_W}) A_{Meta_B} \quad (12)$$

E. Protection with backup VNF

For each of the protection schemes discussed before, we also consider the possibility of having a backup VNF for each VNF of the SFC.

1) Only VNF Backup

This protection scheme only provides backup to VNFs of the SFC. In this case, we also include the availability of the VNF backup, as follows:

$$A_{VNF_1} = A_{vnf_1} + (1 - A_{vnf_1}) A_{VNFB_1} \quad (13)$$

with

$$A_{VNFB1} = A_{vnfb_1} A_{INVL_1} \quad (14)$$

2) End-to-End Protection with VNF Backup

This scheme provides End-to-End (E2E) protection with additional backup VNFs to each VNF along the SFC. The overall availability is given by

$$A_{SFC} = A_{SFC_W} + (1 - A_{SFC_W}) A_{SFC_B} \quad (15)$$

where the availability on the working path is nothing else than the availability of the unprotected scenario with VNF backup

$$A_{VNF_1} = A_{vnf_1} + (1 - A_{vnf_1}) A_{VNFB_1} \quad (16)$$

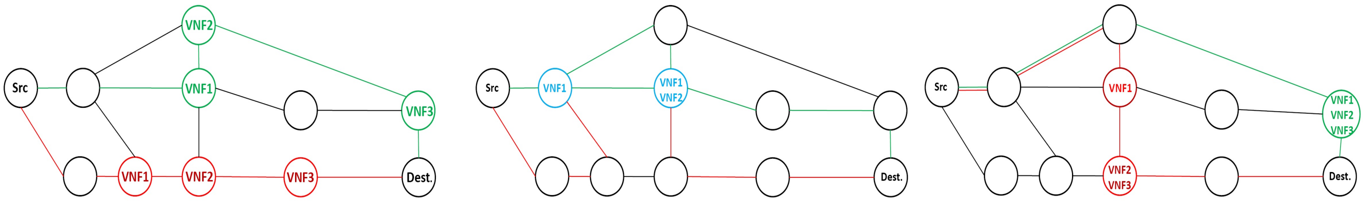


Fig. 2. End-to-End Protection (left); Virtual-Link Protection (middle); Virtual-Node Protection (right).

3) Virtual-Link Protection with VNF Backup

This strategy is the same as Virtual-Link protection, with an additional VNF as backup in the same node. Availability is then given by:

$$P\{E_{SFC}\} = P\{E_{vnt_1} \cap E_{vnt_2} \cap E_{vnt_3} ... \cap E_{vnt_f}\} \quad (17)$$

where each corresponding "Terminal" is equal to

$$P\{E_{vnt_1}\} = P\{E_{vnt_{1_W}} \cup (\overline{E_{vnt_{1_{WP}}}} \cap E_{vntp_{1_B}})\} \quad (18)$$

In this case, the $E_{vnt_{1_W}}$ function is the one with the INVL:

$$E_{vnt_{1_W}} = E_{vnf_1} \cup (\overline{E_{vnf_1}} \cap E_{VNFB_1}) \quad (19)$$

Finally, the total availability is given by the following expression:

$$P\{E_{vnt_1}\} = P\{(E_{vnf_1} \cup (\overline{E_{vnf_1}} \cap E_{VNFB_1}) \cup (\overline{E_{vnt_{1_{WP}}}} \cap E_{vntp_{1_B}})\} \quad (20)$$

4) Virtual-Node Protection with VNF Backup

This strategy consists of protecting every VNF with a backup VNF co-located in the same VN. The communication part of the SFC is not protected. The objective of this strategy is to protect against (typically software) failures of the VNF. Then, the availability is given by the following expressions:

$$P\{E_{SFC}\} = P\{E_{Meta_1} \cap E_{Meta_2} \cap E_{Meta_3} ... \cap E_{Dest_f}\} \quad (21)$$

$$P\{E_{Meta}\} = P\{E_{Meta_W} \cup (\overline{E_{Meta_{WP}}} \cap E_{Meta_p})\} \quad (22)$$

where:

$$P\{E_{Meta_{WP}}\} = P\{E_{Meta_W} \cap (E_{vnf} \cup (\overline{E_{vnf}} \cap E_{VNFB})) \quad (23)$$

IV. Numerical Results

This section discusses numerical results. In our analysis, we have focused on SFC availability requirements in range 0.999 to 0.999999, while for network components VNF availability is in range 0.999 to 0.99999 and link availability is in range 0.9999 to 0.999999.

We have considered SFCs consisting of 3 and 6 VNFs as in [12], each deployed on a separate node. The cost considered for each protection scheme has been varied, based on the number and type of network components (nodes and links) utilized. Specifically, we consider different cost for utilizing network components, which we set at 1, 10 and 2 for using a switching node, a link and a computing node, respectively. The overall costs of protection schemes are reported in Tab. II.

TABLE II
Cost List

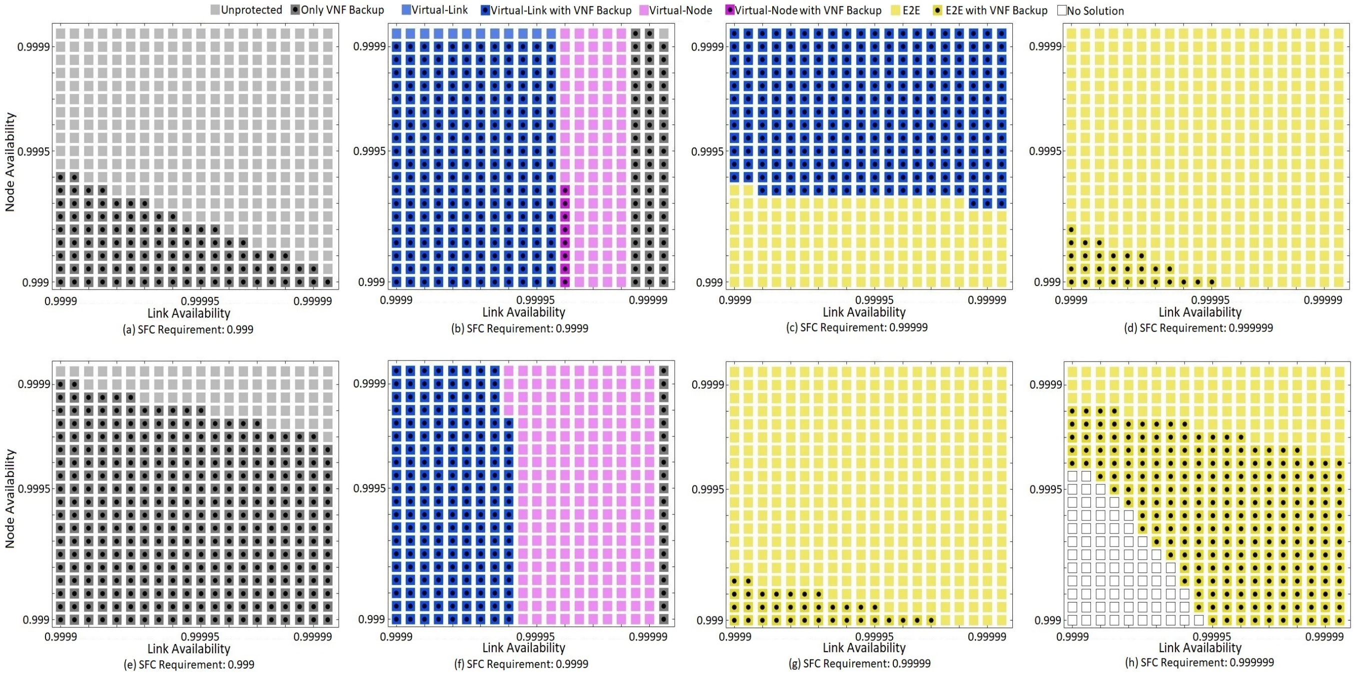
No. of Nodes Protection	3	6
Unprotected	45	82
E2E Protection	90	164
Virtual-Link Protection	87	155
Virtual-Node Protection	68	151
Unprotected with VNF Backup	47	88
E2E with VNF Backup	94	176
Virtual-Link with VNF Backup	89	161
Virtual-Node with VNF Backup	72	163

A selection of results of our numerical analysis is plotted in Fig. 3. Graphs identify the lowest-cost protection scheme that meets availability requirement of SFC, for different SFC availability requirements, varying VNF availability and link availability, in the case of a SFC consisting of 3 VNF nodes (Figs. 3 (a) to (d)) and in the case of a SFC consisting of 6 VNF nodes (Figs. 3 (e) to (h)).

The graphs can be read as follows. The y-axis reports values of node availability (ranging from 0.999 to 0.9999) and the x-axis reports values of link availability (ranging from 0.9999 to 0.99999). Each square refers to a combinations of node availability and link availability and indicates the lowest-cost protection strategy that guarantees meeting the availability requirement of the SFC.

First, we consider the case of a 3-node SFC. For SFC availability requirement equal to 0.999 (Fig. 3a), we see that *Unprotected* and *Only VNF Backup* are sufficient. Specifically, *Only VNF Backup* is required when node availability is lower than 0.9995 and, in some cases, even when link availability is around 0.99999, while *Unprotected* is sufficient for all the remaining cases.

For SFC availability requirement equal to 0.9999 (Fig. 3b), several protection schemes come into play. When link availability is relatively low (below 0.99995), *Virtual link with VNF backup* and *Virtual Link* protection schemes are sufficient, as it is enough to protect virtual links between VNFs to meet SFC availability requirement. For higher link availability (independently of node availability), *Virtual node*



(a) SFC Requirement: 0.999

(b) SFC Requirement: 0.9999

(c) SFC Requirement: 0.99999

(d) SFC Requirement: 0.999999

(e) SFC Requirement: 0.999

(f) SFC Requirement: 0.9999

(g) SFC Requirement: 0.99999

(h) SFC Requirement: 0.999999

Fig. 3. Results reporting lowest-cost protection scheme that meets SFC availability requirement for varying availability requirement (from three 9s to six 9s) for 3 VNs (upper row) and six VNs (lower row) and for varying node and link availability.

with VNF backup and *Virtual node* protection schemes meet SFC availability requirement, while, for relatively high virtual link availability values, it is sufficient to employ *Only VNF Backup* protection scheme. This shows that *E2E* protection can be avoided for 0.9999 availability requirement, and in most cases, depending on availability of network components, relatively lower cost protection schemes can be applied, always considering a 3-node SFC.

For SFC availability requirement equal to 0.99999 (Fig. 3c), *E2E* protection is only required in some cases when node availability is lower than 0.9995. For the remaining cases, *virtual link with VNF backup*, which costs slightly less than *E2E*, meets the availability requirement of the SFC. This further shows that even for a stringent availability requirement, *E2E* protection can be avoided. Finally, for SFC availability requirement equal to 0.99999, results show that *E2E* or E2E with VNF backup are the only possible protection schemes where in particular, E2E with VNF backup is required for relatively low availability of node and links.

Then, we consider the case of a 6-node SFC to analyze how results change for a longer SFC. Note that costs of protection schemes differ with respect to the case of an SFC with 3 nodes and a protection scheme, that has lower cost relatively to other protection schemes for a 3-node SFC, may yield a relatively higher cost in case of 6-node SFC.

For the case of SFC availability of 0.999 and 0.9999 (Figs. 3e and 3f), the least-cost protection schemes guaranteeing the availability target are the same as in the case 3-node SFC, however protection schemes of relatively higher cost are required in more cases. This is because more the replication of more network components is required to meet availability requirements. Yet, in both cases, *E2E* protection can be avoided and, hence, less network-resource occupation can be achieved.

For SFC availability of 0.99999, unlike in the case of 3-node SFC, we see that only *E2E* and *E2E with VNF Backup* protection schemes can guarantee the availility target, due to the fact that more network components need to be protected. In particular, for some specific combinations of low availability of nodes and links, *E2E with VNF Backup* (i.e., the protection scheme with highest cost) is required to meet SFC availability requirement.

Finally, for the most stringent SFC availability target of 0.999999, results show that for combinations of relatively low availability values of nodes and links, no protection scheme can reach availability requirements (white boxes in the figure). For higher availability values, *E2E* and *E2E with VNF Backup* protection schemes are applicable, with *E2E with VNF Backup* dominating most of the cases, due to the very stringent availability requirements.

V. Conclusion

In this paper, we have modelled the cost of several protection schemes for SFCs and, for each scheme, we have evaluated analytically the SFC availability using closed-form equations from classical reliability theory. Compared to existing studies in this area, we consider the availability of the communication links (not only of the VNFs) and we incorporate cost considerations in our analysis. Our numerical results focused on finding the lowest-cost protection scheme that meets availability requirement of SFCs, considering different availability values of network components and length of service function chain. Results are computed for different values of SFC availability requirements, SFC length (number of NFV nodes), availability of network elements, reporting the lowest-cost protection strategy that meets a target availability requirement. Obtained results show that protection schemes with highest cost, as E2E protection, are not always necessary to meet a 0.9999 availability requirement. In most cases, depending on network element availability, lower cost protection schemes can be adopted, to be selected depending on the elements that are characterized by lower availability.

References

[1] Chowdhury, Mosharaf Kabir, et al., “Virtual network embedding with coordinated node and link mapping,” IEEE INFOCOM Conference, 2009.

[2] Askari, Leila, et al. “Virtual-network-function placement for dynamic service chaining in metro-area networks,” International Conference on Optical Network Design and Modeling (ONDM), 2018.

[3] Callegati, Franco, et al. “Dynamic chaining of virtual network functions in cloud-based edge networks,” 1st IEEE Conference on Network Softwarization (NetSoft), 2015.

[4] Liu, Junjie, et al, “On dynamic service function chain deployment and readjustment,” in IEEE Transactions on Network and Service Management, Vol. 14, no. 3, pp. 543-553, 2017.

[5] Allybokus, Zaid, et al., “Virtual function placement for service chaining with partial orders and anti-affinity rules,” in Networks, vol 71, no. 2, pp 97-106, 2018.

[6] Jiang, Huihui, et al., “Availability-aware survivable virtual network embedding in optical datacenter networks,” in IEEE/OSA Journal of Optical Communications and Networking, vol. 7, no. 12, pp 1160-1171, 2015.

[7] Wang, Meng, et al., “Availability-aware service chain composition and mapping in NFV-enabled networks,” IEEE International Conference on Web Services (ICWS), 2019.

[8] Hmaity, Ali, et al., “Virtual network function placement for resilient service chain provisioning,” 8th International Workshop on Resilient Networks Design and Modeling (RNDM), 2016.

[9] Arci, Daniele, et al., “Availability models for protection techniques in WDM networks,” 4th International Workshop on Design of Reliable Communication Networks (DRCN), 2003.

[10] Askari, Leila, et al., “A techno-economic evaluation of VNF placement strategies in optical metro networks,” 4th International Conference on Computing, Communications and Security (ICCCS), 2019.

[11] Alahmad, Yanal, and Anjali, Agarwal, “VNF placement strategy for availability and reliability of network services in NFV,” 6th International Conference on Software Defined Systems (SDS), 2019.

[12] Askari, Leila, et al., “Protection Strategies for Dynamic VNF Placement and Service Chaining,” 2021 International Conference on Computer Communications and Networks (ICCCN), 2021.

[13] Chanclou, Philippe, et al., “The place of optical access to meet the challenges of 5G latency and reliability,” Optical Networking and Communication Conference & Exhibition (OFC), (2020).

[14] Neto, L. Anet, et al., “Enabling technologies and innovations for 5G-oriented optical access networks,” Broadband Access Communication Technologies XV. Vol. 11711. SPIE, 2021.

[15] Ayoub, Omran, et al., “Service Chaining in Filterless Optical Metro-Aggregation Networks,” Asia Communications and Photonics Conference, Optica Publishing Group, 2021.

[16] Ruiz, L., et al. ”Comparison of different protection schemes in the design of VNF-mapping with VNF resiliency,” 22nd International Conference on Transparent Optical Networks (ICTON), 2020.

[17] Kong, Jian, et al. ”Guaranteed-availability network function virtualization with network protection and VNF replication,” IEEE Global Communications Conference (GlobeCom), 2017.