Robust Virtual Network Function Provisioning Under Random Failures on Network Function Enabled Nodes

Tachun Lin
Department of Computer Science and Information Systems
Bradley University, Peoria, IL USA
Email: djlin@bradley.edu

Zhili Zhou
United Airlines
Chicago, IL USA
Email: zhili.zhou@united.com

Abstract—Network function virtualization enables on-demand network function (NF) deployment providing agile and dynamic network services. Early works on NF focused on its provisioning, design, and management with quality metrics – NF-service availability and reliability under system failure(s). To facilitate flexible NF service recovery and migration with high reliability against random NF-enabled node failures, with a known NF resource pool, we first introduce a new NF service evaluation metric to quantify the minimum reliability among all requested NFs for all end-to-end demands – a tight lower bound on individual NF’s service reliability among all requests. We then study the robust virtual network function (VNF) provisioning problem where only a limited number of VNF instances may be instantiated while maximizing the proposed evaluation metric. We present exact solution approach which guarantees the minimum reliability of all NF service to be in the range [76%, 94%] even when physical nodes may fail with a very high (50%) probability.

Index Terms—Robustness of network function service, network function virtualization, cross-layer network, Quality-of-Service (QoS).

I. INTRODUCTION

Network function virtualization (NFV) allows network functions (NFs) to be realized as on-demand services without deploying costly proprietary networking hardware, which serves as a building block supporting the key features and development of the fifth Generation (5G) telecommunication systems [1]. Through NFV, physical resources can be allocated or reallocated to instances of VNFs, thus it not only provides more flexibility but also shortens the enabling time of new NF services [2]. NF recovery and migration are the major approaches to guarantee the continuity, resilience, and security of NF services [3–5]. When VNF manager/orchestrator [6] cannot reach NF instances, it initiates the fail-over to other available NF instances and automatically recovers NF services and/or instantiates new VNFs [7]. Meanwhile, dynamic and flexible VNF migration also reduces power consumption and the burden on hardware capacity [3].

In NFV, end-to-end demands are realized through network flows passing through a series of network functions (with or without a specific sequence) deployed onto NF-enabled containers/servers in the physical infrastructure [8], [9]. “Service function chaining” (SFC) and “non-chained” NFs are commonly used to denote the deployment of the required NFs with or without a sequential order, respectively [10]. Though SFC poses more strict requirements on resource provisioning and dependencies among NF instances, in this paper we unveil a robust and generalized VNF provisioning approach which is suitable for both non-chained NFs and SFC. We adopt cross-layer network structures to model the logical demands and the underlying network infrastructure.

Network infrastructure failures and system attacks threaten the performance, resilience, and security of NF services [11] and the underlying telecommunication networks, which motivate extensively studies on network reliability under a single-layer network setting [12–14]. The reliability analysis of a cross-layer network [15–17] imposes more challenges than its single-layer counterpart as it evaluates the probability of the logical network to remain connected after (random) physical link failure(s). Cutsets [18–20] and spanning trees [21–23] are the two typical network structures utilized to estimate network reliability.

ETSI [24], a leading investigation agency of NFV, defines the end-to-end NF reliability/availability to be the probability that NF “components have not failed after a time period” with known NF-instance locations. With this definition, if both NF deployment and network component failure probabilities are known a priori, Casazza et al. [25] showed that the best (fractional) NF assignment guaranteeing high availability can be determined through backup VNFs. In [26], a neural network-based machine learning algorithm was proposed which exploits the information of VNF forwarding graph to predict future resource requirement. With known physical link failures and backup resources, Soualah et al. [27] used Weibull distribution in formulating the meantime between failures and proposed a decision tree approach targeting the full recovery of NF services. Ding et al. [28] determined backup VNFs and their deployment considering the resources and reliability of the physical nodes. Bijwe et al. [29] assigned priorities to important VNFs to reduce NF service downtime.

While the above studies provide valuable insights from different aspects of NF services, they cannot be used to quantify system capability and reliability to support NF recovery and migration [30–32] as well as seamless NFV state transitions [33], [34] under component failure(s). In this paper we propose an evaluation metric on robust VNF provisioning which measures the minimum NF reliability (the tight lower
bound of NF reliability) among all required NFs in the NF resource pool for all demands, which is also suitable for the SFC requests. Our goal is to provide VNF managers/orchestrators a way to evaluate the strategies to instantiate VNFs on available NF-enabled nodes (NF resource pools) based on the information of the physical infrastructure and resource utilization. We further study a robust VNF provisioning problem with the objective to maximizing the proposed evaluation metric with limited number of VNFs [28] instantiated from the NF resource pools to reduce redundancy and improve physical resources utilization.

Our contributions are summarized as follows. (1) Given NF resource pools and failure probabilities on physical nodes, we propose an evaluation metric on the reliability among all NFs, which is suitable for both non-chained NFs and SFC. (2) We present the robust NF provisioning problem which determines the minimum NF instantiation with a guaranteed low bound on the reliability of NF services. (3) We propose an exact solution approach based on mixed-integer programming. (4) Our work may be extended to the robust design and analysis of interdependent systems (integrated with facility location).

The rest of the paper is organized as follows. In Section II, we introduce the evaluation metric and define the robust VNF provisioning problem, followed by its solution approach in Section III. The experimental settings and simulation results are presented in Section IV. We conclude our work in this paper and future research directions in Section V.

II. NOTATIONS AND PROBLEM DESCRIPTION

In this section, we first provide the general notations used in the discussions. We then propose the robust NF-service evaluation metric and study the robust VNF provisioning problem which minimizes the number of instantiated VNFs while maximizing the robust NF-service evaluation metric. Let \( G_P = (V_P, A_P) \) or \( G_P = (V_P, E_P) \) be the physical infrastructure with node set \( V_P \) and arc set \( A_P \) or edge set \( E_P \). Let \( G_L = (V_L, E_L) \) denote the logical network composed of end-to-end service requests \( D = \{d_{st}\} \), \( s, t \in V_L, (s, t) \in E_L \), and the required NF set \( F \). Let node set \( V_P^f \) denote a physical resource pool for NF \( f \) (candidate physical nodes to deploy \( f \)) and \( V_P^f = \cup_{f \in F} V_P^f \) be the NF-enabled node set. Each of the NF-enabled nodes is with failure probability \( \rho_i \), \( i \in V_P^f \) and \( 0 \leq \rho_i \leq 1 \).

We assume that the NF requests \( d_{st}, s, t \in V_L \) are known a priori. Let \( d_{st} \) be a tuple \([s, t, \sigma_{st}, F_{st}]\), where \( \sigma_{st} \) indicates whether the request is with SFC or not; if yes, \( \sigma_{st} = 1 \), otherwise \( \sigma_{st} = 0 \). We let \( \mathcal{P} \) and \( \overline{\mathcal{P}} \) be the undirected and directed path sets in the physical network. A demand \( d_{st} \) with NF requests is fulfilled if it is routed through path \( p_{st} \in \mathcal{P} \) visiting all required NFs in the sequence defined in SFC when \( \sigma_{st} = 1 \), or otherwise routed through undirected path \( \eta_{st} \in \overline{\mathcal{P}} \) visiting all required NFs with \( \sigma_{st} = 0 \). To simplify the notation, we let \( \mathcal{P} \) represent the path set containing all undirected and directed paths of all NF requests.

Notations and parameters are summarized in Table I.

A. Robust NF Service Evaluation Metric

Our robust NF-service evaluation metric is based on the following observations.

**Observation 1:** Given an NF-enabled node pool \( V_P^F \) and requests \( D = \{d_{st}\} \), where request \( d_{st} \) is realized through a path \( \eta_{st} \). \( d_{st} \) cannot be fulfilled if and only if \( V_P^f \cap \eta_{st} = \emptyset \), \( f \in F_{st} \).

Observation 1 is derived from the fact that \( d_{st} \) can only be fulfilled if and only if (all) the required NFs are deployed onto physical node(s) in its selected path \( \eta_{st} \).

If \( F_{st} = \{f\} \) and \( \eta_{st} \) is chosen for all \( d_{st} \), \( \text{Prob}(d_{st}) \), the probability of \( d_{st} \) being fulfilled, is then \( (1 - \prod_{i \in V_P^f \cap \eta_{st}} \rho_i) \).

We now consider a more generalized setting where demands are with single or multiple NFs and their routings \( \eta_{st} \) are not selected (but with candidates \( \mathcal{P}_{st} \)).

**Definition 1:** Given NF-enabled node pool \( V_P^F \), the robust NF-service evaluation metric, denoted as \( \mathcal{R}\mathcal{P}(d_{st}) \), is

\[
\mathcal{R}\mathcal{P}(d_{st}) = \min_{f \in F_{st}, \eta_{st} \in \mathcal{P}_{st}} \left[ 1 - \prod_{i \in \Gamma(f) \cap \eta_{st}} \rho_i \right].
\]

Note here that \( d_{st} \) with multiple non-chained NF requests is fulfilled if and only if all required NFs are satisfied. Thus, the robust evaluation metric \( \mathcal{R}\mathcal{P}(d_{st}) \) is determined by the worst best-case scenario among all requested NFs realized through

---

### Table I

<table>
<thead>
<tr>
<th>Notation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( G_P = (V_P, A_P) )</td>
<td>Physical infrastructure with ( V_P ) as its node set and ( A_P ) as its arc and edge sets, respectively</td>
</tr>
<tr>
<td>( G_L = (V_L, E_L) )</td>
<td>Logical network with ( V_L, E_L ) as its node and edge sets, respectively</td>
</tr>
<tr>
<td>( i )</td>
<td>An NF-enabled node</td>
</tr>
<tr>
<td>( s, t, (s, t) )</td>
<td>A request ( d_{st} )</td>
</tr>
<tr>
<td>( P, \eta_{st}, P_{st} )</td>
<td>( P ) represents the path sets in ( G_P ), where ( \eta_{st} \in P ) denote the directed and undirected paths, respectively. ( P_{st} = {\eta_{st}, P_{st}} )</td>
</tr>
<tr>
<td>( F, F_{st}, f )</td>
<td>( F ) represents all the NFs; ( F_{st} ) denotes the required NFs for end-to-end demand ( (s, t) ); and ( f ) in ( F ) is a network function</td>
</tr>
<tr>
<td>( V_P^f )</td>
<td>A set of all NF-enabled nodes, ( V_P^f \subseteq V_P )</td>
</tr>
<tr>
<td>( \Gamma(F), n'_{st} )</td>
<td>( \Gamma(F) = {{f, i, n'<em>{st}} : f \in F, i \in V_P^f} ) is the deployment of NF instances, where ( n'</em>{st} ) is the instances of ( f ) deployed onto ( i )</td>
</tr>
</tbody>
</table>

---

Notation and parameters are summarized in Table I.
the best-known paths in $P_{st}$. Hence, $RP$ provides an estimated low bound of NF-service reliability for all demands.

Different from the non-chained NF requests, SFC requests are fulfilled only when all required NFs are served in a specified sequence. Without loss of generality, we assume that (1) the same NF request will not be fulfilled more than once on different NF-enabled nodes, and (2) each NF-enabled node will not carry out multiple NF requests in SFC.

Definition 2: Given NF-enabled node pool $V^F_p$, the robust NF evaluation metric of SFC request $d_{st}$ is

$$RP(d_{st}) = \min_{f \in F_{st}, p_{st} \in \mathcal{P}_{st}} \max_{\rho_i \in \Gamma_f \cap \mathcal{P}_{st}} \left[ 1 - \prod_{i \in \Gamma(f)} \rho_i \right] / |F_{st}|!.$$ 

Since demands with SFC request are fulfilled only when all requested NFs are deployed onto $p_{st}$ and visited in a predefined sequence, there is only one valid case out of $|F_{st}|!$ permutations. $RP(d_{st})$ is then determined by the worst best-case scenario among all requested NFs realized through the best-known paths in $P_{st}$ (with the highest probability to survive).

Considering multiple NF requests in a given NFVI (NFV infrastructure) and managed by the same NFV MANO (NFV management and organization), we define the robust NF-service evaluation metric among all NF requests as follows.

Definition 3: Given $G_P, G_L$, a set of NFs $F$, NF-enabled node pool $V^F_p$ and node failure probability $\rho_i, i \in V^F_p$, $RP(V^F_p) = \min_{d_{st}, \in D} RP(d_{st})$.

B. Illustrations: NF Service Reliability vs. Robust NF Service Evaluation Metric

We evaluate the robust NF-service evaluation metric via an instance illustrated in Fig. 1 and present its difference to the NF-service reliability defined in [24]. In this example, two demands with NF requests $d_{12}$ and $d_{34}$ are considered. Demand $d_{12}$ requires SFC $f_1 \rightarrow f_2$ and $d_{34}$ requires non-chained NFs $\{f_1, f_2\}$. NF-enabled nodes, their supported NFs, and their failure probabilities are labeled in Fig. 1(b). Candidate physical nodes to enable/deploy $f_i$’s are in set $V^1_p = \{1, 3, 4, 5\}$, and those for $f_2$’s are in $V^2_p = \{2, 3, 5, 6\}$. $d_{12}$ is routed through a directed path $\{(1, 5), (5, 2)\}$, and $d_{34}$ is routed through an undirected path $\{(4, 6), (6, 3)\}$. Based on the assumptions given in the previous section, the robust NF-service evaluation metric $RP(\{d_{12}, d_{34}\}) = \min\{1 - 0.1, 1 - 0.2, (1 - 0.2 \times 0.1)/2, (1 - 0.1 \times 0.2)/2\} = 0.49$.

Different from $RP(\{d_{12}, d_{34}\})$, NF-service reliability of $d_{12}$ is

$$1 - \text{Prob}(f_1, f_2 \text{ both failed}) - \text{Prob}(\text{only } f_2 \text{ failed})$$
$$- \text{Prob}(\text{only } f_1 \text{ failed}) - \text{Prob}(f_1, f_2 \text{ fulfilled not in-order})$$
$$= 1 - 0.1 \times 0.2 \times 0.1 - 0.9 \times 0.2 \times 0.1 - 0.9 \times 0.2 \times 0.1 - 0$$
$$= 0.962.$$ 

The NF-service reliability of $d_{34}$ is

$$1 - \text{Prob}(f_1, f_2 \text{ both failed}) - \text{Prob}(\text{only } f_2 \text{ failed})$$
$$- \text{Prob}(\text{only } f_1 \text{ failed}) - \text{Prob}(f_1, f_2 \text{ fulfilled not in-order})$$
$$= 1 - 0.2 \times 0.1 \times 0.1 - 0.2 \times 0.1 \times 0.9 - 0.2 \times 0.9 \times 0.1$$
$$= 0.962.$$ 

The examples above show that NF-service reliability is measured when the deployment of NF instances and routings are determined. In contrast, since the robust NF-service evaluation metric already evaluates the minimum NF reliability, the routings selected and the deployment of non-chained NFs or SFC would always be better than or at least equal to the metric. In other words, the robust NF-service evaluation metric provides a tight lower bound for each NF’s reliability.

This instance also shows that with the limitation imposed on the NF-enabled nodes, the selection of NF-enabled nodes also impact the robust NF-service evaluation metric. Hence, in the following section, we study the robust NF provisioning problem which aims at maximizing our proposed NF-service evaluation metric via NF-enabled node selection for NF request realization.

C. Robust NF Provisioning

When taking the failures of NF-enabled nodes into consideration, we now define the robust VNFM provisioning problem. Given $G_P, G_L, D$, and $V^F_p$, $d_{st} \in D, s, t \in V_L$, is mapped onto a directed path $p_{st} \in P$ for SFC request, or undirected path $\eta_{st} \in P$ for NF request. We would like to determine a limited number of NF-enabled nodes to support each required NF and guarantee that demands are routed through their required NFs. This problem considers both non-chained NF and SFC requests.
In this section, we discuss the failure probability of NF and SFC requests. Hence, we have the following proposition.

Proposition 1: \( 1 - \mathcal{R}(V_p|f) = \mathcal{F}(V_p|f) \), with \( d_{st} \in D \) and \( f \in F \).

Proposition 1 derives directly from Definitions 1 and 2, which also holds for SFC requests. Hence, we have the following conclusion.

Theorem 1: \( \max_{V_p \in F_p} \mathcal{R}(V_p|f) = \min_{V_p \in F_p} \mathcal{F}(V_p|f) \).

Based on Proposition 1 and Theorem 1, we have the following equations.

\[
\mathcal{F}(d_{st}) = \max_{f \in F_{st}} \min_{\eta_{st} \in \mathcal{P}_{st}} \left[ \prod_{i \in \Gamma(f) \cap \eta_{st}} \rho_i \right] 
\]  \tag{1}

\[
\mathcal{F}(d_{st}) = \max_{f \in F_{st}} \min_{\eta_{st} \in \mathcal{P}_{st}} \left[ \prod_{i \in \Gamma(f) \cap \eta_{st}} \rho_i \right] / |F_{st}| \]  \tag{2}

\[
\mathcal{F}(V_p|f) = \max_{d_{st} \in D} \mathcal{F}(d_{st}) \]  \tag{3}

In the next section, we demonstrate that solving the robust VNF provisioning via minimizing NF failure evaluation metric provides a linearization of non-linear equations. We then propose the solution approach accordingly.

III. SOLUTION APPROACH

In this section, we demonstrate how to utilize \( \mathcal{F}(V_p|f) \) to formulate robust VNF provisioning problem and propose the mixed-integer programming solution approach for the problem. The variables and parameters used in this section are presented in Table II.

A. Formulations for NF Request

We now present the mathematical formulations for the maximal reliable NF deployment problem based on the NF service failure probability. We first turn the non-linear objective \( \min_{V_p} \max_{d_{st} \in D} \min_{\eta_{st} \in \mathcal{P}_{st}} \Pi_{i \in V_p \cap \eta_{st}} \rho_i \) into its linearized counterpart

\[
\min \max_{V_p} \min_{d_{st} \in D} \min_{\eta_{st} \in \mathcal{P}_{st}} \sum_{i \in V_p \cap \eta_{st}} \ln(\rho_i) 
\]  \tag{4}

by applying the \( \ln(\cdot) \) function.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( N_f )</td>
<td>The number limitation for NF deployed locations with ( f \in F )</td>
</tr>
<tr>
<td>( \rho_i )</td>
<td>The failure probability of physical node ( i ) with ( i \in V_p )</td>
</tr>
<tr>
<td>( \delta_{st} )</td>
<td>A binary indicator showing whether physical node ( i ) is on path ( \eta_{st} ) or not, ( \eta_{st} \in \mathcal{P}<em>{st} ), ( (s,t) \in E_L ); if yes, ( \delta</em>{st} = 1 ), otherwise ( \delta_{st} = 0 )</td>
</tr>
<tr>
<td>( \gamma_{st} )</td>
<td>A binary indicator showing whether ( f ) is requested by ( d_{st} ) or not; if yes, ( \gamma_{st} = 1 ), otherwise ( \gamma_{st} = 0 )</td>
</tr>
<tr>
<td>( M )</td>
<td>A very large number</td>
</tr>
</tbody>
</table>

TABLE II

PARAMETERS AND VARIABLES

With Theorem 1, the formulation presented below is the robust NF evaluation metric value of NF request with (4) as the objective:

\[
\min_{\lambda,x,y,z,\xi,h} \lambda \sum_{i \in V_p} h_i \leq N_f, f \in F \]  \tag{5}

\[
\lambda \geq \xi_{st}, f \in F, d_{st} \in D \]  \tag{6}

\[
\xi_{st} = \sum_{i \in V_p} \ln(\rho_i) y_{st,i}, f \in F, d_{st} \in D \]  \tag{7}

\[
y_{st,i} \geq z_{st,i} + \delta_{\eta_{st},x_{\eta_{st}}} + \gamma_{st} - 2, f \in F, d_{st} \in D, \eta_{st} \in \mathcal{P}_{st}, i \in V_p \]  \tag{8}

\[
y_{st,i} \leq z_{st,i}, f \in F, i \in V_p \]  \tag{9}

\[
y_{st,i} \leq \delta_{\eta_{st},x_{\eta_{st}},f}, f \in F, i \in V_p, d_{st} \in D, \eta_{st} \in \mathcal{P}_{st} \]  \tag{10}

\[
y_{st,i} \leq g_{st}, f \in F, i \in V_p, d_{st} \in D \]  \tag{11}

\[
h_i \geq z_{st,i}, f \in F, i \in V_p \]  \tag{12}

\[
\sum_{\eta_{st} \in \mathcal{P}_{st}} x_{\eta_{st}} = 1, d_{st} \in D \]  \tag{13}

\[
\lambda, \xi_{st} \geq 0, z_{st,i}, y_{st,i}, h_i, x_{\eta_{st}} \in \{0,1\}, \eta_{st} \in \mathcal{P}_{st}, (s,t) \in E_L, f \in F, d_{st} \in D, i \in V_p \]  \tag{14}

Constraint (5) enforces the upper bound for the number of
nodes deployed with NFs. Constraint (6) records the value of NF failure evaluation metric (linearized) among all demands for all NFs. Constraint (7) captures the robust NF failure evaluation metric value (linearized, i.e., \( \ln(\rho_i) \)) as in constraint (4)) of demand \( d_{st} \in D \) and \( f \in F \). Based on Definition 1, constraint (8) determines whether \( f \) is deployed onto physical node \( i \) for demand \( d_{st} \in D \), where (i) \( z^f_i = 1 \) when \( f \) is deployed onto physical node \( i \); (ii) \( \delta^i_{st} = 1 \) when node \( i \) deployed with an NF is on a selected path \( \eta_{st} \) for \( d_{st} \); and (iii) \( \gamma^i_{st} = 1 \) when \( d_{st} \) requires NF \( f \). Constraints (9) – (11) force variable \( \eta^f_{st} \) to be 0 when any of the (i) to (iii) above is not satisfied. Constraint (12) indicates whether physical node \( i \) is deployed with any NFs. Constraint (13) selects a single physical route for demand \( d_{st} \in D \). Constraint (14) provides feasible regions for all variables.

Note here that the variable \( \lambda \) in constraint (6) records the value of the robust NF failure evaluation metric achieved by NF request through \( \xi^f_{st} \). As the objective of the reformulation is to find the minimum \( \lambda \), it also encourages evaluation metric value \( \xi^f_{st} \) to be minimized. Therefore, the above reformulation solves the maximal reliable NF deployment problem.

We next present the formulation for SFC service reliability.

**B. Formulations for SFC Service Reliability**

Different from the non-chained NF failure probability, the SFC failure probability is

\[
1 - \max_{d_{st} \in D} \max_{\eta_{st} \in P_{st}} \left[ 1 - \prod_{i \in \Gamma(f) \cap \rho_{st}} \rho_i \right] / |F_{st}|!
\]

where \( d_{st} \in D \).

**Proposition 2:** For requests with SFC, we have

\[
1 - \min_{d_{st} \in D} \max_{f \in F_s} \min_{\rho_{st} \in P_{st}} \left[ \prod_{i \in \Gamma(f) \cap \rho_{st}} \rho_i \right] / |F_{st}|!
\]

where \( F_s \) represents the requested NFs of \( d_{st} \). We introduce here an auxiliary variable \( \omega_{st} \) which indicates whether \( d_{st} \in D \) is selected as the \( d^*_st \). By replacing routings from undirected to directed path set (i.e., \( \eta_{st} \rightarrow \rho_{st} \)) in constraints (8), (10), (13), (14), we present the formulation for the robust SFC provisioning as follows.

\[
\begin{align*}
\min_{\lambda, \xi, \omega, \beta, y, x, z} & \quad \lambda \\
\text{s.t.} & \quad \lambda \geq \omega_{st}, \\
& \quad d_{st} \in D \quad (15) \\
& \quad \omega_{st} \geq \xi^f_{st} - \ln |F_{st}|!, \quad f \in F, d_{st} \in D \quad (16) \\
& \quad \sum_{d_{st} \in D} \beta_{st} = 1 \quad (17) \\
& \quad \lambda \leq \omega_{st} + M(1 - \beta_{st}), \quad d_{st} \in D \quad (18) \\
& \quad \lambda \geq \omega_{st} + M(\beta_{st} - 1), \quad d_{st} \in D \quad (19) \\
& \quad \omega_{st} \geq 0, \beta_{st} \in \{0, 1\}, d_{st} \in D \quad (20) \\
\end{align*}
\]

Constraints (5) and (7)–(14) are created. The first testing cases for the maximal reliable NF deployment problem have (i) NSF as the physical network, (ii) demands with up to three randomly assigned NF requests, (iii) a given limitation on the number of NSF deployed locations, and (iv) random node failure probability. The proposed setting is to verify that when the number of NF locations decreases,
whether the NF service reliability also goes down correspond-
ing. Meanwhile, when the node failure probability increases,
whether the NF service reliability also decreases.

The second testing cases have (i) a fixed NF service reliabil-
ity (90%), and (ii) random physical node failure probability.
The purpose of the setting is again to evaluate that with a
fixed NF service reliability, when the node failure probability
increases, whether extra NF-deployed nodes are required to
fulfill the requirement of the service level.

We report the simulation results with the two sets of testing
cases in the following section.

B. Simulation Results

The simulation results for the maximal NF reliable deploy-
ment problem are presented in Fig. 3. The three lines in
blue, red, and green colors represent the testing cases with
40%, 50%, and 60% of NF-enabled physical nodes. The x-
axis represents the physical node failure probability (in mean
value) and the y-axis denotes the NF service reliability (in
percentage). Each plotted node/dot in the figure presents the
average NF service reliability for all testing samples. With up
to 50% failure probability of the NF-enabled nodes, the NF re-
liability reaches 75%. When the number of NF-enabled nodes
increases, the NF reliability increases to 87.5%, and 93.7%,
respectively. We confirm our analysis that with the limitation
on the number of NF-enabled nodes, the NF service reliability
increases when physical node failure probability decreases.
Also, given the same physical node failure probabilities, we
observe that when the number of NF-enabled nodes (in terms
of the mean value) decreases, the reliability of the NF service
decreases as well.

Figure 4 illustrates the number of NFs deployed to reach
the required level of the NF service reliability (based on the
maximal number of NF-enabled nodes in the testing cases)
with single NF and multiple NFs (in our testing cases, three
required NFs) in each demand. To reach the fixed (90%)
NF service reliability, the number of physical nodes deployed
with NFs is only doubled when the number of required NFs
for each demand goes from one to three even with high
failure probability (10 – 50%) on physical nodes. The figure
demonstrates a clear pattern between the number of nodes
deployed with NFs and the NF service reliability.

In the simulation results, we observe that the NF service
reliability is higher with more physical nodes deployed with
the required NFs, and obviously, a lower average node failure
probability leads to a higher NF service reliability under the
failure(s) of physical nodes. The observations on these simula-
tions are as expected and demonstrate the relationship between
the number of NF-deployed nodes (cost-related restriction) and
NF service reliability (service level).

V. CONCLUSION AND FUTURE WORKS

We studied the reliable NF deployment problems under
random physical node failure in this paper. We proposed an
evaluation metric, the NF service reliability, to quantify and
indicate the probability of the required NFs being fulfilled
for the end-to-end demands. Utilizing this evaluation metric,
we studied the maximal NF reliable deployment problem
and proposed the exact solution approaches to solve the
problem. We designed and conducted simulations to confirm
our analysis on the reliable NF deployment approaches.

In the further research, we would also like to consider
physical node capacity and NF deployment costs, and evaluate
the costs to introduce more NF-enabled nodes in the physical
network. We also like to investigate the scenarios of shared risk
group failure(s) and physical link failure(s) and their impacts
on NF service reliability. Last, but not least, another research
direction is to relax the assumptions on independent node
failures, the correlations among NF-enabled node failures, and
study their impacts on NF service reliability.

REFERENCES

R. Boultaba, “Network function virtualization: State-of-the-art and re-
search challenges,” IEEE Communications Surveys & Tutorials, vol. 18,
service function chain routing and virtual function network instance
migration in network function virtualization architectures,” IEEE/ACM
updates for virtualized network function migration,” IEEE Transactions
on Services Computing, 2017.


