2022 IEEE 22nd International Conference on Software Quality, Reliability and Security (QRS)

Availability and Cost aware Multi-Domain Service Deployment Optimization

Chuangchuang Zhang, Yanming Liu, Hongyong Yang, Yihao Li*, and Shuning Zhang
School of Information and Electrical Engineering, Ludong University, Yantai 264025, China
zhangcc_neu@163.com, liuyanming19981016@163.com, Yihao.li@ldu.edu.cn
*corresponding author

Abstract—**Network Function Virtualization (NFV) achieves flexible provisioning of network services by using Service Function Chain (SFC) composed of a set of Virtual Network Functions (VNFs). However, complex multi-domain networks pose serious challenges to multi-domain service deployment with availability guarantee. In this paper, we study the availability and cost aware multi-domain service deployment optimization problem. We formulate a multi-objective optimization model with the aim to minimize resource consumption cost and operating cost, while guaranteeing availability by jointly considering VNF failures and server failures, as well as cross-domain deployment operating cost. Then, we design a VNF backup based multi-domain SFC deployment algorithm to reduce resource consumption cost and operating cost. The evaluation results demonstrate that our proposed algorithm can achieve lower resource consumption cost and operating cost than comparison algorithms.**

Keywords- Network function virtualization, Service function chain, Availability guarantee, multi-domain networks

I. INTRODUCTION

As an emerging network architecture, Network Function Virtualization (NFV) [1] decouples traditional network functions from dedicated hardware devices and implements them in software, called Virtual Network Functions (VNFs), which can be flexibly deployed in common-off-the–shelf servers. Each network service can be represented as one Service Function Chain (SFC) in NFV, which is composed of a set of ordered VNFs [2]. By leveraging NFV technology, capital expenditures and operating expenses in service provisioning can be saved significantly.

With the rapid development of network technologies, such as 5G and IoTs, future networks are required to provide higher quality network services, and service requirements from users, e.g, low latency and high reliability, should be satisfied [3]. Service availability is of paramount important to fulfill Service Level Agreement (SLA) and guarantee efficient service provisioning [4]. Compared with traditional applications, five 9s or six 9s (i.e., 99.999% or 99.9999%) need to be guaranteed for network services. However, complex multi-domain network environments impose great challenges in efficient service provisioning with availability guarantee.

There exist a large number of heterogeneous network devices in multi-domain network scenarios [5, 6]. Moreover, multiple service providers and network operators corporate to support multi-domain service provisioning. The resource usage costs in different network domains are generally different [7]. Software failures and hardware failures have impacts on multi-domain SFC deployment with availability guarantee. Different SFC deployment strategies can result in different service availability and deployment cost.

To improve service availability, current proposals adopt backup and redundancy methods in SFC deployment process [8, 9]. Although backup or redundancy strategies can be used to reduce the impact of failures, it incurs more resource consumption. If primary and backup VNFs are cross-domain mapped in different server nodes, it incurs additional operating cost, and degrades user’s service Quality of Experience (QoE). How to deploy the SFC and backup VNFs across multiple network domains becomes a critical issue for multi-domain service provisioning with service availability guarantee. Thus, it is necessary to design an efficient multi-domain SFC deployment optimization mechanism to guarantee service availability while reducing resource consumption and operating cost.

To this end, in this paper, we investigate the multi-domain SFC deployment optimization with service availability guarantee and SFC cross-domain deployment cost optimization. We firstly formulate the availability and cost aware multi-domain SFC deployment problem as a multi-objective optimization model to minimize resource consumption cost and operating cost. Then, a multi-domain SFC deployment optimization approach is presented to deploy SFCs and backup VNFs in multi-domain networks. The main contributions are summarized as follows.

(1) We formulate the availability and cost aware multi-domain SFC deployment problem as a multi-objective optimization model with the target of resource consumption cost minimization and operating cost minimization by jointly considering VNF failures and server failures.

(2) We propose an availability and cost aware multi-domain SFC deployment algorithm by using backup and redundancy strategies to reduce resource consumption cost and operating cost.

(3) We conduct extensive simulations for performance evaluation. Evaluation results demonstrate that our proposed deployment algorithm can have smaller resource

2693-9177/22/$31.00 ©2022 IEEE
DOI 10.1109/QRS57517.2022.00051

consumption cost and operating cost than comparison mechanisms.

The rest of this paper is organized as follows. Section 2 discusses the related work, and system model and problem formulation are described in Section 3. The proposal solution is presented in Section 4. Section 5 evaluates the performance of our proposed solution. Finally, Section 6 concludes this study.

II. RELATED WORK

In this section, we discuss related work about service availability optimization of multi-domain SFC deployment.

To improve service availability, Li et al [10] designed an availability aware VNF deployment and backup solution. Sharing mechanism of redundancy and multi-tenancy technology are used to improve resource utilization efficiency, and an availability calculation algorithm for shared redundancy based backup scheme is designed to calculate the modified availability of a VNF. Alahmad et al [11] formulated the SFC deployment problem as an Integer Linear Programming (ILP) optimization model with the target of service reliability maximization, and proposed a VNF placement strategy to improve service availability and reliability. Mandal et al [12] analyzed the service availability of the SFC placement in three cases, including being deployed in multiple host nodes, single host node and mixed-mode by jointly considering VNF failures and host node failures, and further analyzed the service reliability of SFC placement in three cases. Wang et al [13] studied the parallelized SFC placement problem in data center networks to guarantee service availability and optimize resource, and proposed availability and traffic aware parallelized SFC placement model. The backup model provides backup sub-SFCs for working sub-SFCs to improve service availability and a hybrid placement algorithm is designed.

Zhang et al [14] presented a sub-chain-enabled coordinated protection model to achieve availability aware SFC provisioning. The protection model provides protection for each sub-chain to satisfy the SFC availability in a cost-efficient method. An optimization model is formulated to minimize deployment cost, and a heuristic approach is designed to solve it. Xu et al [15] designed a high availability SFC placement approach in data center networks to improve service availability. The proposed approach uses offsite redundant VNF instances to avoid physical machines failures and guarantee service availability by considering data link reliability. A heuristic limited search optimization algorithm is presented to place the SFC. Kang et al [16] designed a primary and backup SFC placement model to avoid service interruptions and guarantee service availability, and formulated an ILP optimization model to maximize the minimum number of continuously available time slots in SFCs. Zhang et al [17] formulated the availability-guaranteed service function chain (SFC) provisioning problem as an optimization model to minimize deployment cost, and proposed a sub-chain-enabled coordinated protection model to configure sub-chains for each SFC and provide proper protection for each sub-chain to guarantee service availability.

To improve service availability while reducing latency, Yala et al [18] formulated the latency and availability driven VNF placement problem in MEC-NFV environment as a multi-objective optimization model with the target of access latency minimization and service availability maximization, and proposed a genetic algorithm based heuristic VNF placement approach to solve it. Similarly, De Simone et al [19] proposed a latency driven availability multi-tenant service chains assessment model to evaluate the availability and latency, and design a modified version of multi-dimensional universal generating function techniques.

Mauro et al [20] designed a novel high availability service chain management framework to automatically build service chains with satisfying service availability requirements in a minimal cost manner. Yin et al [21] studied the SFC placement problem in MEC-NFV scenarios, and designed a backup model to improve the SFC availability. A dynamic programming based SFC placement algorithm is proposed to optimize resource cost. To avoid backup resource inefficiency and guarantee service availability, Araújo et al [22] proposed an optional backup with shared path and shared function SFC provisioning method. The proposed method uses the optional backup resources to assign the VNF instances and physical links only when needed.

To guarantee high availability and reduce energy consumption, Abdelaal et al [23] formulated the SFC deployment problem in cloud computing as an ILP optimization model with the target of link utilization, convergence time and energy consumption minimization, and designed a redundant VNF forwarding graph heuristic deployment algorithm to find a trade-off between availability and scalability. Santos et al [24] designed a Reinforcement Learning (RL) based SFC placement solution in large scale networks to achieve the balance between service availability, SFC placement cost and energy consumption. The proposed placement algorithm uses RL to determine the suitable candidate nodes and uses the redundancy strategy to satisfy service availability requirement. And further, Santos et al [25] proposed an availability and energy aware SFC placement approach by leveraging RL to support dynamic SFC placement. Two policy aware RL algorithms are used to improve the SFC availability and energy consumption. Mail et al [26] studied the energy efficiency optimization problem with service availability guarantee, and designed an energy efficient SFC placement framework to reduce energy consumption and guarantee service availability.

The above research efforts try to optimize the SFC deployment to improve service availability by considering different optimization objectives, such as resource consumption, energy consumption, latency, reliability.

However, few studies pay attentions to the SFC deployment with availability guarantee in the multi-domain network scenarios. Different from existing works, in this paper, we focus on the multi-domain SFC deployment problem with service availability guarantee to jointly reduce resource consumption cost and SFC deployment operating cost.

III. SYSTEM MODEL AND PROBLEM FORMULATION

In this part, we make a formulation about availability and cost aware multi-domain SFC deployment optimization problem. Table 1 lists the basic notations used in this paper.

Table 1. Notations

Symbol	Definition
Network	
G	The substrate multi-domain network
V	The set of physical nodes
E	The set of physical links
G^i	The i-th network domain
$v_{i,j}^s$	The j-th server in G^i
$e_{m,n}^{i,j}$	The physical link between nodes $v_{m,n}$ and $v_{i,j}$
$A_s^{i,j}$	The availability of server $v_{i,j}^s$
SFC	
SR_i	The i-th service request
(src_i, dst_i)	Source and destination node of SR_i
sf_i	The SFC of SR_i
bw_i	The bandwidth requirement of SR_i
$sf_{i,j}$	The j-th VNF in SFC sf_i
$r_{i,j}^{cpu}$	The CPU requirement of $sf_{i,j}$
$r_{i,j}^{mem}$	The memory requirement of $sf_{i,j}$
$r_{i,j}^{str}$	The storage requirement of $sf_{i,j}$
$e_{i,j,k}$	The virtual link between $sf_{i,j}$ and $sf_{i,k}$
A_{sf}^i	The availability demand of SFC sf_i
$A_f^{i,j}$	The availability of $sf_{i,j}$ VNF instance
Resource	
$C_{i,j}^{cpu}$	The CPU capacity of server $v_{i,j}^s$
$C_{i,j}^{mem}$	The memory capacity of server $v_{i,j}^s$
$C_{i,j}^{str}$	The storage capacity of server $v_{i,j}^s$
$bw_{m,n}^{i,j}$	The bandwidth capacity of link $e_{m,n}^{i,j}$
$Pr_{i,j}^{cpu}$	The unit prices of CPU resource
$Pr_{i,j}^{mem}$	The unit prices of memory resource
$Pr_{i,j}^{str}$	The unit prices of storage resource
$Pr_{i,j,u,v}^{bw}$	The unit prices of bandwidth resource
Variable	
$x_{i,j}^{m,n}$	1 if $sf_{i,j}$ is placed in $v_{i,j}^s$, 0 otherwise
$y_{i,j,k}^{m,n,u,v}$	1 if $e_{i,j,k}$ is mapped in $e_{m,n}^{u,v}$, 0 otherwise
$z_{i,j,k}^{s,d}$	1 if $e_{i,j,k}$ is mapped in p_s^d, 0 otherwise

In this paper, we make several assumptions for simplicity as follows. Each VNF instance fails independently, and for each VNF, the primary and backup instances have the same availability and resource demands.

A. System Model

(1) Network model

The substrate multi-domain network is modeled as an undirected graph $G = (V, E)$, where V is the set of physical nodes and E is the set of physical links. In this paper, we consider that physical nodes consist of switch nodes (v^w) and server nodes (v^s). $G^i = (V^i, E^i)$ is used to indicate the i-th network domain in multi-domain network G. Each server $v_{i,j}^s$ in network domain G^i has a finite amount of resource, such as CPU, memory. For simplicity, in this paper, we only consider CPU, memory and storage resource attribute for each server. The resource capacity of each server $v_{i,j}^s$is denoted by $C_{i,j}^{cpu}$, $C_{i,j}^{mem}$ and $C_{i,j}^{str}$, respectively. Each physical link $e_{m,n}^{i,j}$ between physical nodes $v_{m,n}$ and $v_{i,j}$ is associated with certain bandwidth resource, denoted by $C_{m,n,i,j}^{bw}$. The unit prices of CPU, memory, storage and bandwidth resource are denoted by $Pr_{i,j}^{cpu}$, $Pr_{i,j}^{mem}$ $Pr_{i,j}^{str}$ and $Pr_{i,j,u,v}^{bw}$, respectively.

(2) service request

In NFV, VNFs are chained in a predefined order as an SFC to provide a specific network service for users. Let $SR_i = (src_i, dst_i, sf_i, r_i^{bw})$ be i-th service request, where src_i and dst_i denote source and destination of service request, sf_i denotes the SFC, r_i^{bw} denotes bandwidth requirement of service request.

For each SFC, it can be modeled as $sf_i = (sf_{i,1}, sf_{i,2}, \dots, sf_{i,j}, \dots)$. For each VNF, we define $r_{i,j}^{cpu}$, $r_{i,j}^{mem}$ and $r_{i,j}^{str}$ to represent its CPU, memory and storage resource requirement, respectively. The virtual link between VNF nodes $sf_{i,j}$ and $sf_{i,k}$ is denoted by $e_{i,j,k}$.

(3) Availability model

The availability refers to the probability that a system can work normally within a period of time. In this paper, we only consider the availability of server and VNF. The availability of server and VNF can be calculated as follows.

$$A = \frac{uptime}{uptime+downtime} = \frac{MTBF}{MTBF+MTTR} \quad (1)$$

Where, $uptime$ is the time that a server or VNF is in operation regularly, i.e., Mean Time Between Failures ($MTBF$), $downtime$ is the time that server or VNF is out of service, i.e., Mean Time To Repair ($MTTR$).

We denote the availability of each server $v_{i,j}^s$ and VNF $f_{i,j}$ by $A_s^{i,j}$ and $A_f^{i,j}$, respectively. Thus, the availability of SFC sf_i can be calculated as:

$$A_i = \prod_j A_f^{i,j} \quad (2)$$

B. Problem Formulation

The SFC deployment in multi-domain networks consists of two stage, i.e., VNF placement and virtual link mapping. To facilitate the problem formulation of availability and cost aware multi-domain service deployment, we define some binary decision variables as follows.

The VNF placement decision variable $x_{i,j}^{m,n}$ equals to 1 if VNF $sf_{i,j}$ is successfully placed in server node $v_{i,j}^s$, otherwise, $x_{i,j}^{m,n}$ equals to 0.

$$x_{i,j}^{m,n} = \begin{cases} 1, & if\ sf_{i,j}\ is\ placed\ in\ v_{m,n}^s \\ 0, & otherwise \end{cases} \quad (3)$$

The link map decision variable $y_{i,j,k}^{m,n,u,v}$ is defined to identify virtual link map relationship. $y_{i,j,k}^{m,n,u,v} = 1$ if virtual link $e_{i,j,k}$ is successfully mapped in physical link $e_{m,n}^{u,v}$, otherwise, 0.

$$y_{i,j,k}^{m,n,u,v} = \begin{cases} 1, & if\ e_{i,j,k}\ is\ mapped\ in\ e_{m,n}^{u,v} \\ 0, & otherwise \end{cases} \quad (4)$$

Further, we define a binary variable $z_{i,j,k}^{s,d}$ to describe whether virtual link $e_{i,j,k}$ is mapped in physical path p_s^d from source node s to destination node d, as follows.

$$z_{i,j,k}^{s,d} = \begin{cases} 1, & if\ e_{i,j,k}\ is\ mapped\ in\ z_{i,j,k}^{s,d} \\ 0, & otherwise \end{cases} \quad (5)$$

We formulate the availability and cost aware multi-domain SFC deployment problem as a multi-objective optimization model. Our optimization objectives are expressed as follows.

$$min \quad rc = \sum_{i,j,m,n} x_{i,j}^{m,n} \cdot \left(r_{i,j}^{cpu} \cdot Pr_{m,n}^{cpu} + r_{i,j}^{mem} \cdot Pr_{m,n}^{mem} + r_{i,j}^{str} \cdot Pr_{m,n}^{str}\right) + \sum_{sf_i, e_{m,n}^{u,v}, p_s^d} y_{i,j,k}^{m,n,u,v} \cdot z_{i,j,k}^{s,d} \cdot Pr_{m,n,u,v}^{bw} \quad (6)$$

$$min \quad pc = \sum_{i,j,m,n} \mu \cdot (\emptyset_1(sf_i) + H(sf_i) \cdot \emptyset_2(sf_i)) \quad (7)$$

Where rc denotes the total resource consumption cost, pc denotes the operating cost caused by cross-domain transmission of data flow, $\mu \in (0,1)$ denotes the operating cost weight coefficient, $\emptyset_1(sf_i)$ denotes the total number of network domains occupied by the SFC sf_i when it is successfully deployed for the first time, $\emptyset_2(sf_i)$ denotes the total number of network domains occupied by the SFC sf_i when it is successfully re-deployed again due to VNF or server failures, $H(sf_i)$ denotes the indicator function, $H(sf_i) = 1$ represents the SFC is re-deployed again due to failures, otherwise, $H(sf_i) = 0$.

Regarding each VNF, it should be placed in a single server rather than multiple servers. Thus, the constraint of VNF placement is:

$$\sum_{m,n} x_{i,j}^{m,n} = 1 \quad (8)$$

Similarly, the virtual link map constraints are formulated as:

$$\sum_{m,n,u,v} y_{i,j,k}^{m,n,u,v} = 1 \quad (9)$$

$$\sum_{s,d} z_{i,j,k}^{s,d} = 1 \quad (10)$$

With respect to resource consumption, we take into account CPU, memory, storage and bandwidth resources. The resource requirements of all the VNFs placed in same server cannot exceed the resource capacity of this server. The resource constraints are described as follows.

$$\sum_{i,j} x_{i,j}^{m,n} \cdot r_{i,j}^{cpu} \le C_{m,n}^{cpu} \quad (11)$$

$$\sum_{i,j} x_{i,j}^{m,n} \cdot r_{i,j}^{mem} \le C_{m,n}^{mem} \quad (12)$$

$$\sum_{i,j} x_{i,j}^{m,n} \cdot r_{i,j}^{str} \le C_{m,n}^{str} \quad (13)$$

$$\sum_{sf_i} y_{i,j,k}^{m,n,u,v} \cdot r_i^{bw} \le C_{m,n,u,v}^{bw} \quad (14)$$

The availability of SFC must be ensured by guaranteeing VNF and server are available, as follows.

$$\prod_j \left(x_{i,j}^{m,n} \cdot A_f^{i,j} \cdot A_s^{m,n}\right) \le A_i \quad (15)$$

In addition, all binary decision variables should obey the integer constraints.

IV. THE AVAILABILITY AND COST AWARE MULTI-DOMAIN SFC DEPLOYMENT APPROACH

A. Algorithm design

To solve the above problem, we propose an availability and cost aware multi-domain SFC deployment algorithm (ACSP) by considering backup and redundancy strategies.

In our solution, to reduce resource consumption cost, we try to deploy the VNFs in the SFC in the server nodes with small resource usage price and use the backup resources when and only when the primary SFC cannot satisfy the service availability. Moreover, to reduce the operating cost, we try to deploy the VNFs in the SFC in same network domains as possible.

The specific workflow of our proposed algorithm is as follows, described in Algorithm 1.

At first, we determine the candidate server nodes which have higher availability not smaller than service availability requirement of the SFC.

Then, the following procedure is executed until all the VNFs in the SFC are successfully deployed or the SFC request is rejected.

(1) For the first VNF, we select the servers from candidate servers which have abundant resource capacity and are located in same network domain with source node of the SFC.
(2) If the appropriate servers are not fund in the network domain where source node of the SFC are located, we continue to search appropriate servers from candidate server nodes in other network domains.
(3) If all candidate servers cannot satisfy the resource demand of VNF, the SFC request is rejected.
(4) If the appropriate servers are fund, we deploy the VNF in the server (labeled by *S1*) with minimum resource usage price, and determine the physical link mapping between source node of the SFC and this server node by using the shortest path algorithm.
(5) For the second VNF in the SFC, similar to the deployment process of first VNF, we select its candidate servers from candidate servers which can satisfy resource demand of second VNF and are located in same network domain with server node *S1*.
(6) If candidate servers can be fund, we deploy the second VNF in the optimal server (labeled by *S2*) with minimum resource usage price, and determine the physical link mapping between server nodes *S1* and *S2* by using the shortest path algorithm.
(7) We repeat the above processes to deploy the remaining VNFs in the SFC by repeating the above operations.

Next, when and only when the deployed SFC cannot offer efficient network service due to VNF failures or server failures, we adopt the backup resource and redundancy strategies.

(1) If the VNF fails, we re-deploy a VNF instance in the corresponding server.

(2) If the server (labeled by S) cannot efficiently work due to failures, we re-determine an optimal server for the VNF (labeled by *vf*).

(2.1) We check whether the last hop or next hop server of server S can satisfy the service availability and resource demand of VNF *vf.* If the conditions are met, we re-deploy VNF in the last hop or next hop server node, and calculate the shortest physical link by using Dijkstra algorithm.

(2.2) If the last hop and next hop server cannot efficiently satisfy service availability requirement and resource demand of VNF, we search the appropriate servers from all the candidate servers with satisfying service availability and VNF resource demand.

(2.3) If all the servers cannot satisfy the requirement, the SFC request is rejected. Otherwise, we re-deploy the VNF in the server with minimum resource usage price, and map the physical link by determining the shortest path.

Algorithm 1: ACSP algorithm

Input: The multi-domain network G;
Service request SR
Output: Deployment solution dps

```
Begin
Determine candidate servers S with satisfying $A_{sf}^{i}$;
For each VNF $sf_{i,j}$ in SFC $sf_i$
    IF $sf_{i,j}$ is the first VNF
        $pre = src_i$;
    EndIF
    S*=SelectCandidateServer(S, $r_{i,j}^{cpu}$,$r_{i,j}^{mem}$,$r_{i,j}^{str}$, $pre$);
    IF S*=Null
        S*=SelectCandidateServer(S, $r_{i,j}^{cpu}$,$r_{i,j}^{mem}$,$r_{i,j}^{str}$);
    EndIF
    IF S*!=Null
        S1=DeployVNF(S*,$sf_{i,j}$);
        MapPhysicalLink($pre$,S1);
        $pre = S1$;
    Else
        Reject $sf_i$;
    EndIF
EndFor
IF VNF vf fails
    Re-deploy a VNF instance in the server;
EndIF
IF server S fails
    IF last hop or next hop of server S meets conditions
        Re-deploy VNF vf in last hop or next hop server;
    Else
        S*=SelectCandidateServer(S, vf);
        IF S*!=Null
            S1=DeployVNF(S*, vf);
        Else
            Reject $sf_i$;
        EndIF
    EndIF
EndIF
Record the SFC placement solution in $dps$;
End
```

B. *Complexity analysis*

We assume that the multi-domain network has N physical nodes, M physical links, K server nodes, N_G network domains. The total amount of service requests is N_{sfc}, the maximum length of SFC is L_{sfc}. So, the time complexity of the proposed ACSP algorithm can be expressed as $T = O\left(N_{sfc} \cdot L_{sfc} \cdot M \cdot log^{N}\right)$.

V. PERFERMANCE EVALUATION

In this section, we conduct performance evaluations for ACSP scheme, and analyze simulation results.

A. *Simulation Settings*

The simulation programs are run on win7 PC, Inter Core(TM), 2.93GHz CPU, 4GRAM, using MATLAB 2015b platform.

To simulate real multi-domain network scenarios, we select CERNET2 (C) and Interoute (I) networks as test topologies, and randomly divide them into four and ten network domains, respectively. CERNET2 consists of 20 network nodes and 22 physical links, and Interoute is comprised of 110 network nodes and 148 physical links. For each physical node in both scenarios, it is selected as server with the probability 0.5. The availability of each physical node obeys the uniform distribution of (0.99, 0.999).

In the simulations, we use "unit" to quantity the resource capacity, resource demand and resource usage cost, respectively. In CERNET2, resource capacity of each server is a random integer between 100 and 200, and bandwidth capacity of each physical link is a number distributed uniformly between 200 and 400. In Interoute, resource capacity of each server obeys the uniform distribution between 10 and 50, and bandwidth capacity of each physical link obeys the uniform distribution between 50 and 100. The unit prices of all resources in both scenarios are set to obey the uniform distribution between 1 and 5, respectively.

Each SFC and its source and destination nodes are generated randomly. The bandwidth demand of each SFC obeys uniform distribution of (1, 10). The length of each SFC is an integer distributed uniformly between 2 and 5. Five types of VNFs are considered. The resource demands of different types of VNFs obey the uniform distribution between 1 and 5. The availability of each type of VNF randomly varies from 0.99 to 0.999. The exponential distribution parameter γ of service request arrival rate is set as 5, 10 and 15, respectively, and the Poisson distribution parameter of service lifetime is set as 20. Additionally, we set $\mu = 1.0$.

The random deployment and backup based multi-domain SFC deployment algorithm (RSP) and the first-fit deployment and backup based multi-domain SFC deployment algorithm (FFSP) are used compared with our proposed ACSP algorithm. RSP algorithm randomly selects appropriate servers with satisfying resource and availability demands to deploy VNFs, and FFSP algorithm gives the priority to the first servers with satisfying resource and availability demands to deploy VNFs in a first-fit manner.

To comprehensively evaluate the performance of our proposed deployment algorithm, the following indexes are used for comparisons.

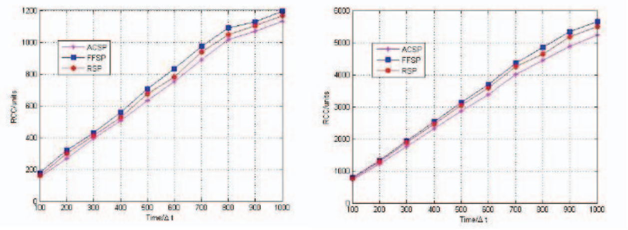
(1) Resource Consumption Cost (RCC): It refers to the total resource consumption usage cost caused by the SFC deployment, including primary and backup resources.

(2) OPerating Cost (OPC): It refers to the operating cost caused by the SFC cross-domain deployment.

(3) Time Overhead (TO): It refers to the time that the multi-domain SFC deployment algorithm takes to deploy a set of SFC requests.

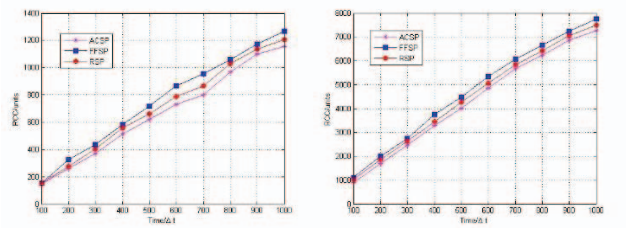
B. Simulation Results

(1) RCC



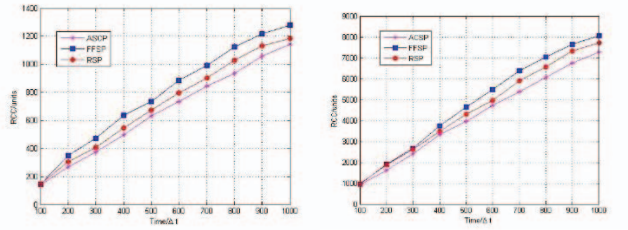
(a) RCC (C, $\gamma = 5$) (b) RCC (I, $\gamma = 5$)

Figure 1. RCC comparison



(a) RCC (C, $\gamma = 10$) (b) RCC (I, $\gamma = 10$)

Figure 2. RCC comparison



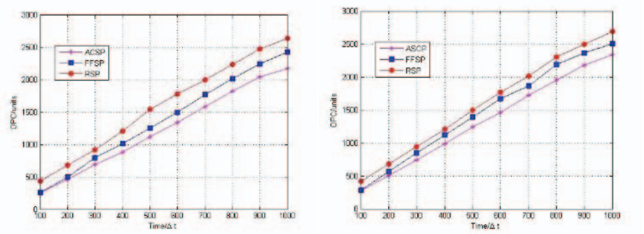
(a) RCC (C, $\gamma = 15$) (b) RCC (I, $\gamma = 15$)

Figure 3. RCC comparison

We can observe from Figs. 1-3 that with the increasing of time periods, the overall resource consumption cost gradually becomes high. This is because more SFCs are deployed in multi-domain networks with the increase in time periods, and more resources are consumed. Compared with RSP and FFSP algorithms, the proposed ACSP algorithm has lower resource consumption cost. This is because ACSP algorithm tries to deploy the VNFs in the servers with small resource usage price, and the other two algorithms do not consider resource usage price selection. Moreover, we observe that with the increasing of parameter λ, the resource consumption cost of three algorithms becomes big. The reasons are explained as follows. The parameter increase means that more service requests are generated and deployed, and more SFCs are deployed in multi-domain networks.

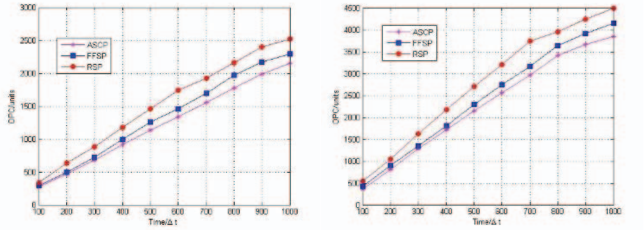
(2) OPC

From Figs. 4-6, it can be observed that ACSP algorithm has smaller operating cost than RSP and FFSP algorithms. The detailed reasons are as follows. Compared with RSP and FFSP algorithm, ACSP algorithm tries to deploy the VNFs in the SFC in same network domains as possible. The SFC uses fewer network domains by using ACSP algorithm than the other two algorithms. Moreover, with the increasing of time periods, the operating cost of three deployment algorithm becomes big. This is because with the increase in time periods, more service requests are generated and more SFCs are deployed in multi-domain networks.



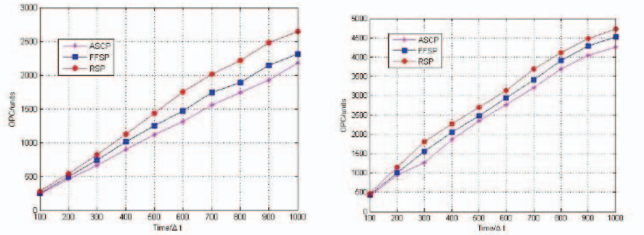
(a) OPC (C, $\gamma = 5$) (b) OPC (I, $\gamma = 5$)

Figure 4. OPC comparison



(a) OPC (C, $\gamma = 10$) (b) OPC (I, $\gamma = 10$)

Figure 5. OPC comparison

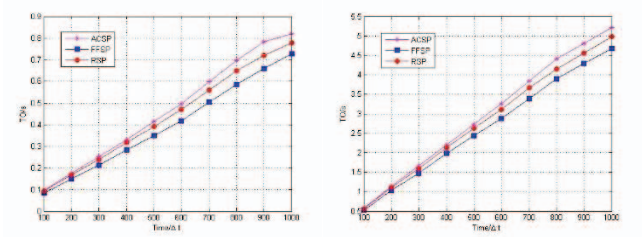


(a) OPC (C, $\gamma = 15$) (b) OPC (I, $\gamma = 15$)

Figure 6. OPC comparison

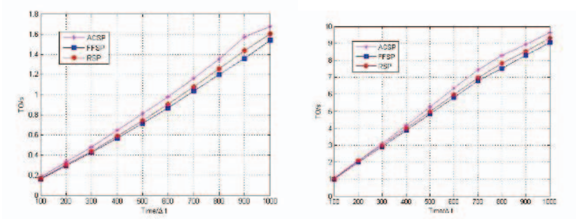
(3) TO

As depicted in Figs. 7-9, we can observe that with the increasing of time periods, the time overhead of three deployment algorithms becomes big. This is because with the increase in time periods, more service requests are generated. The deployment algorithms take more time to deploy the SFCs in multi-domain networks. We can also observe that compared with RSP and FFSP algorithms, ACSP algorithm takes more time for SFC deployment due to involving with candidate server selection, shortest path calculation and backup adjustment process. In addition, it can be seen that compared with CERNET2 network, the deployment algorithms take more time for SFC deployment in Interoute network. This is because Interoute network has more physical nodes and physical link than CERNET2 network.



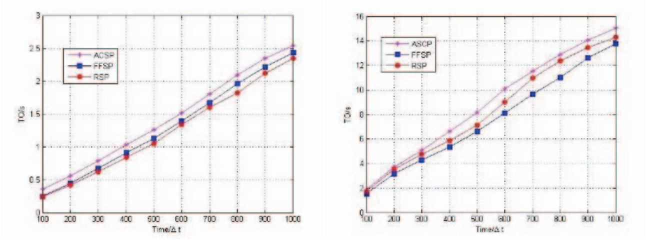
(a) TO (C, $\gamma = 5$) (b) TO (I, $\gamma = 5$)

Figure 7. TO comparison



(a) TO (C, $\gamma = 10$) (b) TO (I, $\gamma = 10$)

Figure 8. TO comparison



(a) TO (C, $\gamma = 15$) (b) TO (I, $\gamma = 15$)

Figure 9. TO comparison

VI. CONCLUSION

In this paper, we study the multi-domain SFC deployment by considering service availability guarantee, resource consumption cost and operating cost. An availability and cost aware multi-domain SFC deployment optimization model is formulated with the target of cross-domain deployment cost minimization while guaranteeing availability of SFC. Availability and cost aware multi-domain SFC deployment algorithm is designed to solve the above problem. Evaluation results show the proposed solution outperforms comparison schemes in terms of service availability and cross-domain deployment cost.

However, in real multi-domain network scenarios, service requests from users are dynamically uncertain, meanwhile network topology and traffic are dynamically changing. In such case, how to efficiently guarantee service availability and reduce cross-domain deployment cost is one critical challenge. In future works, we will investigate the multi-domain service availability guarantee considering the dynamic of service request and network topology while reducing cross-domain deployment cost.

ACKNOWLEDGMENT

This work is supported by Science Foundation of Shandong Province of China under Grant No. ZR2021QF086, National Natural Science Foundation of China under Grant No. 61872073, 61673200, 61603172 and 62072091, Major Basic Research Project of Natural Science Foundation of Shandong Province of China under Grant No. ZR2018ZC0438.

REFERENCES

[1] Yang S, Li F, Trajanovski S, et al. Recent advances of resource allocation in network function virtualization[J]. IEEE Transactions on Parallel and Distributed Systems, 2020, vol. 32, no.2, pp. 295-314.

[2] Arzo S T, Bassoli R, Granelli F, et al. Study of virtual network function placement in 5G cloud radio access network[J]. IEEE Transactions on Network and Service Management, 2020, vol.17, no. 4, pp.2242-2259.

[3] Guo S, Qi Y, Jin Y, et al. Endogenous trusted DRL-based service function chain orchestration for IoT[J]. IEEE Transactions on Computers, 2021, vol.71, no.2, pp.397-406.

[4] Wang L, Mao W, Zhao J, et al. DDQP: a double deep Q-learning approach to online fault-tolerant SFC placement[J]. IEEE Transactions on Network and Service Management, 2021, vol.18, no.1, pp.118-132

[5] Lin R, Yu S, Luo S, et al. Column generation based service function chaining embedding in multi-domain networks[J]. IEEE Transactions on Cloud Computing, 2021.

[6] Sarrigiannis I, Antonopoulos A, Ramantas K, et al. Cost-aware placement and enhanced lifecycle management of service function chains in a multi-domain 5G architecture[J]. IEEE Transactions on Network and Service Management, 2022.

[7] Zhang C, Wang X, Dong A, et al. Energy efficient network service deployment across multiple SDN domains[J]. Computer Communications, 2020, vol.151, pp.449-462.

[8] Niu M, Han Q, Cheng B, et al. HARS: a high-available and resource-saving service function chain placement approach in data center networks[J]. IEEE Transactions on Network and Service Management, 2022.

[9] Yan B, Zhao Y, Yu X, et al. Service function path provisioning with topology aggregation in multi-domain optical networks[J]. IEEE/ACM Transactions on Networking, 2020, vol.28, no.6, pp. 2755-2767

[10] Li D, Hong P, Xue K, et al. Availability aware VNF deployment in datacenter through shared redundancy and multi-tenancy[J]. IEEE Transactions on Network and Service Management, 2019, vol.16, no.4, pp.1651-1664.

[11] Alahmad Y, Agarwal A. VNF placement strategy for availability and reliability of network services in NFV[C]. 2019 Sixth International Conference on Software Defined Systems (SDS). IEEE, 2019, pp.284-289.

[12] Mandal P. Comparison of placement variants of virtual network functions from availability and reliability perspective[J]. IEEE Transactions on Network and Service Management, 2022.

[13] Wang M, Cheng B, Wang S, et al. Availability-and traffic-aware placement of parallelized SFC in data center networks[J]. IEEE

Transactions on Network and Service Management, 2021, vol.18, no.1, pp.182-194.

[14] Zhang Y, He F, Oki E. Availability-aware service chain provisioning with sub-chain-enabled coordinated protection[C]. 2021 IFIP/IEEE International Symposium on Integrated Network Management (IM). IEEE, 2021, pp.638-642.

[15] Xu Z, Han Q, Cheng B, et al. HASP: high availability SFC placement approach in data center network[C]. 2021 IEEE 23rd Int Conf on High Performance Computing & Communications; 7th Int Conf on Data Science & Systems; 19th Int Conf on Smart City; 7th Int Conf on Dependability in Sensor, Cloud & Big Data Systems & Application (HPCC/DSS/SmartCity/DependSys). IEEE, 2021, pp.448-455.

[16] Kang R, He F, Oki E. Virtual network function allocation in service function chains using backups with availability schedule[J]. IEEE Transactions on Network and Service Management, 2021, vol.18, no.4, pp.4294-4310.

[17] Zhang Y, He F, Oki E. Service chain provisioning with sub-chain-enabled coordinated protection to satisfy availability requirements[J]. IEEE Transactions on Network and Service Management, 2021.

[18] Yala L, Frangoudis P A, Ksentini A. Latency and availability driven VNF placement in a MEC-NFV environment[C]. 2018 IEEE Global Communications Conference (GLOBECOM). IEEE, 2018, pp.1-7.

[19] De Simone L, Di Mauro M, Natella R, et al. A latency-driven availability assessment for multi-tenant service chains[J]. IEEE Transactions on Services Computing, 2022.

[20] Di Mauro M, Galatro G, Longo M, et al. HASFC: a MANO-compliant framework for availability management of service chains[J]. IEEE Communications Magazine, 2021, vol.59, no.6,pp. 52-58.

[21] Yin X, Cheng B, Wang M, et al. Availability-aware service function chain placement in mobile edge computing[C]. 2020 IEEE World Congress on Services (SERVICES). IEEE, 2020, pp.69-74.

[22] Araújo I M, Natalino C, Chen H, et al. Availability-guaranteed service function chain provisioning with optional shared backups[C]. 2020 16th International Conference on the Design of Reliable Communication Networks DRCN 2020. IEEE, 2020, pp.1-6.

[23] Abdelaal M A, Ebrahim G A, Anis W R. High availability deployment of virtual network function forwarding graph in cloud computing environments[J]. IEEE Access, 2021, vol.9, pp.53861-53884.

[24] Santos G L, Endo P T, Lynn T, et al. A reinforcement learning-based approach for availability-aware service function chain placement in large-scale networks[J]. Future Generation Computer Systems, 2022.

[25] Santos G L, Lynn T, Kelner J, et al. Availability-aware and energy-aware dynamic SFC placement using reinforcement learning[J]. The Journal of Supercomputing, 2021, vol.77, no.11, pp.12711-12740.

[26] Mai L, Ding Y, Zhang X, et al. Energy efficiency with service availability guarantee for network function virtualization[J]. Future Generation Computer Systems, 2021, vol.119, pp.140-153.