UDC 621.391+004.7

MODEL FOR ALLOCATION OF NETWORK SLICES IN HETEROGENEOUS ENVIRONMENT

Sulima S.V.
National Technical University of Ukraine
“Igor Sikorsky Kyiv Polytechnic Institute”
E-mail: lilthirteen@gmail.com

МОДЕЛЬ РОЗМІЩЕННЯ МЕРЕЖЕВИХ СЛАЙСІВ У ГЕТЕРОГЕННОМУ СЕРЕДОВИЩІ

У статті наводиться опис моделі розміщення мережевих функцій мережевих слайсів у гетерогенному середовищі, в якому наявні виділені апаратні блоки з фіксованим ресурсом обслуговування та віртуалізовані блоки, ресурси яких налаштовуються.

The article describes the network functions allocation model of network slices in the heterogeneous environment, where there are dedicated hardware blocks with fixed service resource and virtualized blocks with configurable resources.

Global mobile data usage is growing exponentially. Recent innovations in the mobile communication technologies and mobile terminals drive the proliferation of different services with a wide range of demands in terms of latency, mobility and reliability among others [1]. Network slicing has the capability of enabling (through the network architecture) future 5G networks that encompasses the required scalability and flexibility characteristics, thus supporting diverse service scenarios and services. A network slice can be broadly defined as an end-to-end logically isolated network that includes 5G devices as well as access, transport, and core network functions [2].

The network slicing concept is enabled by the Network Functions Virtualization (NFV) technology and closely related to a Virtual Network Embedding (VNE) problem that is presented in the current paper. The problem is aimed at determining the placement and amount of resources dedicated to each network function in each function chain, i.e. network slice. The general embedding approach is extended by defining network functions placement model, incorporating the notion of heterogeneous networks containing both physical devices offering services and virtualized services as well as binding network function productivity with the amount of allocated resource is proposed similar to the approach [3]. Also the differentiation of slicing objectives and its peculiarities are added to the model. The proposed enhanced optimization model incorporates the parameters and variables shown in Table I.

Table I – Model parameters

Physical model	
G	the graph $G=(N,E)$ represents substrate network
N	$n \in N$ set of nodes include compute nodes N^c and physical network functions N^p
E	$e=(n_1,n_2) \in E$ is the set of edges within the network
Γ	$\gamma \in \Gamma$ types of resources
C^{γ}_n	γ resource capacity of node n
$CE_{n1,n2}$	bandwidth capacity of physical link $(n_1,n_2) \in E$
$L_{n1,n2}$	transmission delay of physical link $(n_1,n_2) \in E$
NF model	
$\tau_n^{\ r}$	processing delay of a physical NF r on node n
NS model	
SFC	$ch \in SFC$ set of all service chains that must be allocated
R_{ch}	$r_c \in R_{ch}$ is set of logical node of function chain ch
$L_{ch}^{\ max}$	tolerance delay of fuction chain ch
Variables	
$x^{r,ch}_{\ \ n}$	value is 1 when function chain ch node r is allocated on physical node n
$F_{ch,e,\,r1,\,r2}$	value is 1 when request of function chain ch link (r_1,r_2) is mapping to physical link $e=(n_1,n_2)$
$D^{r,ch}_{\ \ \ \gamma}$	integer specifies the amount of type γ resource used by function chain ch node r
$DE_{ch,r1,\,r2}$	integer specifies the amount of bandwidth used by function chain ch link (r_1,r_2)

The optimization problem is formulated as follows:

Objectives:

$$\min_{x_n^{r,ch},F_{ch,e,r_1,r_2},D_\gamma^{r,ch},DE_{ch,r_1,r_2})} (a \cdot \sum_{n\in N} x_n \cdot NC_n + b \cdot \sum_{n\in N}\sum_{\gamma\in\Gamma}\sum_{ch\in SFC}\sum_{r\in R_{ch}} (virt_n \cdot x_n^{r,ch} \cdot D_\gamma^{r,ch} + phys_n{}^r \cdot x_n^{r,ch}) \cdot NC_n^\gamma + c \cdot \sum_{(n_1,n_2)\in E} NC_{(n_1,n_2)}^{band} \cdot \sum_{ch\in SFC}\sum_{(r_1,r_2)\in R_{ch}{}^2} F_{ch,(n_1,n_2),r_1,r_2} \cdot DE_{ch,r_1,r_2} + d \sum_{ch\in SFC} FC_{ch} \cdot (1 - \Phi^{ch})) \tag{1}$$

$$\max_{F_{(n_1,n_2),r_1,r_2},DE_{ch,r_1,r_2})} \sum_{(n_1,n_2)\epsilon E} (CE_{n_1,n_2} - \sum_{ch\in SFC}\sum_{(r_1,r_2)\in R_{ch}{}^2} F_{ch,(n_1,n_2),r_1,r_2} \cdot DE_{ch,r_1,r_2}) \tag{2}$$

Subject to:

$$\forall n \in N: \forall\gamma \in \Gamma: \sum_{ch\in SFC}\sum_{r\in R_{ch}} (virt_n \cdot x_n^{r,ch} \cdot D_\gamma^{r,ch} + phys_n{}^r \cdot x_n^{r,ch}) \le C_n^\gamma \tag{3}$$

$$\forall ch \in SFC: \forall r \in R_{ch}: \sum_{n\in N} x_n^{r,ch} = 1 \tag{4}$$

$$\forall n \in N: \forall ch \in SFC: \forall r \in R_{ch}: x_n^{r,ch} \le suit_n^{r,ch} \tag{5}$$

$$\forall ch \in SFC: \forall (r_1,r_2) \in R_{ch}{}^2: \forall n \in \mathrm{N}: x_n^{r_2,ch} + \sum_{e\in E_n^{out}} F_{ch,e,r_1,r_2} = x_n^{r_1,ch} + \sum_{e\in E_n^{in}} F_{ch,e,r_1,r_2} \tag{6}$$

$$\forall ch \in SFC: \forall (r_1, r_2) \in {R_{ch}}^2: \forall n \in \mathrm{N}: x_n^{r_2,ch} + \sum_{e \in E_n^{out}} F_{ch,e,r_1,r_2} \le 1 \tag{7}$$

$$\forall ch \in SFC: \forall (r_1, r_2) \in {R_{ch}}^2: \forall n \in \mathrm{N}: x_n^{r_1,ch} + \sum_{e \in E_n^{in}} F_{ch,e,r_1,r_2} \le 1 \tag{8}$$

$$\forall (n_1, n_2) \in \mathrm{E}: \sum_{ch \in SFC} \sum_{(r_1,r_2) \in {R_{ch}}^2} F_{ch,(n_1,n_2),r_1,r_2} \cdot DE_{ch,r_1,r_2} \le CE_{n_1,n_2} \tag{9}$$

$$\forall ch \in SFC: \tag{10}$$

$$\sum_{r \in R_{ch}} \sum_{n \in N} x_n^{r,ch} \cdot \left(virt_n \cdot \sum_{\gamma \in \Gamma} \frac{s_\gamma^r}{D_\gamma^{r,ch}} + {phys_n}^r \cdot {\tau_n}^r \right) + \sum_{(r_1,r_2) \in {R_{ch}}^2} \sum_{(n_1,n_2) \in E} F_{ch,(n_1,n_2),r_1,r_2} \cdot L_{n_1,n_2} \le {L_{ch}}^{max}$$

where *a, b, c, d* – weight coefficients;

NC_n, NC^{γ}_n, $NC^{band}_{(n1,n2)}$ – node, resource and bandwidth cost;

FC_{ch} – *ch* service allocation failure cost.

The objectives of the model are to minimize the total cost of the service function chains allocations (1) and leave more bandwidth on each physical link (2). The constraints are: (3) the available capacity constraint, (4) ensures that every network function shall be mapped only once, (5) ensures that network function can be mapped only on suitable nodes, (6) the flow conservation constraint, (7)-(8) to make sure there is no incoming flow in source nodes or outgoing flow in sink nodes, (9) the available bandwidth constraint, (10) the required latency constraint.

Conclusion. The paper considers the mobile network slicing. The model for resource allocation of the system of data centers for determining optimal amount of the resources allocated to network functions in the slice has been presented. The model can be used in the management of deployment of network functions in heterogeneous hardware environment in order to minimize the operator costs and improve quality of experience.

References

1. Sama M. R. Reshaping the mobile core network via function decomposition and network slicing for the 5G Era / M. R. Sama, X. An, Q. Wei, S. Beker. // 2016 IEEE Wireless Communications and Networking Conference. – Doha, Qatar, 2016. – P. 90-96.
2. Yousaf F. Z. Reshaping the mobile core network via function decomposition and network slicing for the 5G Era / F. Z. Yousaf, M. Gramaglia, V. Friderikos, B. Gajic, et al.// 2017 IEEE International Conference on Communications Workshops. – Paris, France, 2017. – P. 1195-1201.
3. Skulysh M. Model for Efficient Allocation of Network Functions in Hybrid Environment / M. Skulysh, L. Globa, S. Sulima // Information and Telecommunication Sciences. — 2016. — № 1. — P. 39–45.