The Static Virtual Machine Placement and Routing Spectrum Assignment for Multi-tenant in Elastic Optical Networks

FangPing You(✉)

Hubei University of Technology, Wuhan, China
youfp@hbut.edu.cn

Abstract. The static routing spectrum assignment is a fundamental problem in typical elastic optical networks, where the traffic demands are given and stationary. However, an elastic optical data center network is shared by multi-tenant, and hence the traffic demands are varying and determined by the virtual machine (VM) placement of multi-tenant under the given VM demands. In this paper, the problem of virtual machine placement and routing spectrum assignment (VMPRSA) is introduced and the static problem is proved to be NP-Hard. Furthermore, an optimal integer linear programming (ILP) model is formulated, with the target of minimizing the spectrum used to serve all the traffic demands driven by multi-tenant VM placement. Since the ILP model cannot scale to the big networks, a two-tier heuristic algorithm framework is proposed, i.e., the first tier is VM placement and the second is routing spectrum assignment (VMP + RSA). The VM placement of tenants can produce traffic demands, it thus is focused and two VM placement algorithms are proposed, namely random placement (RP) algorithm, and residual node capacity priority (RNCP) algorithm. The simulation results indicate that the ILP model provides the optimal solution, and the RNCP algorithm yields the better sub-optimal solution.

Keywords: Optical · Data centers · Virtual machine placement

1 Introduction

1.1 A Subsection Sample

Data center networks, as the network infrastructures and the platforms for deploying and running many applications of today's business in cloud computing, have attracted significant attentions in recent years [1, 2].

To meet the exponential growth of traffic induced by data-intensive applications, many literatures have begun designing novel network architectures employing off-the-shelf commodity switches. Techniques in these proposals include fat tree and random graph. Furthermore, in contrast to electrical switching, optical switching which possesses huge transmission bandwidth, can benefit data center networks. Many novel optical/electrical hybrid or all optical architectures are presented with the aim to improve

W. Hong and G. Kanaparan (Eds.): ICCSE 2023, CCIS 2023, pp. 72–84, 2024.
https://doi.org/10.1007/978-981-97-0730-0_8

network capacity and scalability [3–6]. Moreover, an advanced modulation technology, orthogonal frequency division multiplexing (OFDM) has been efficiently demonstrated and can implement elastic bandwidth assignment in optical networks [7]. Therefore, we expect that it is necessary for future optical data center networks to employ the modulation technology, which have been demonstrated in papers. We refer the optical data center networks based on OFDM as elastic optical data center networks (EODCNs).

As we known, the problem of resource assignment is investigated by few papers in elastic optical networks recently, including routing and spectrum assignment (RSA) [7], traffic grooming, etc. It seems that the approaches of resource assignment in elastic optical networks can be seamlessly transplanted to EODCNs [8]. However, EODCNs can be shared by multi-tenant in the cloud. The virtual machine (VM) placement can generate traffic demands of all-to-all between VMs for a tenant, since a VM of a tenant should exchange data with all other VMs of the tenant to achieve a complete response to the tenant, which hence poses a novel challenge on resource assignment [9]. Note VMs between different tenants have no correlation and thereby no traffic demand emerges.

In EODCNs, the problem of routing and spectrum assignment has been explored [8]. However, the fundamental problem of placing VMs of multi-tenant and assigning routing spectrum resources, is unexplored. The problem considers each traffic demand driven by VM placement under the given tenants VM demands, while the networks should have the enough node capacity to satisfy the VM placement (in terms of CPU/memory). The problem is defined as virtual machine placement and routing spectrum assignment (VMPRSA) problem.

In this paper, the static VMPRSA problem is studied in EODCNs, with the goal of minimizing the utilized spectrum to serve all the traffic demands determined by VM placement of each tenant. To the best of our knowledge, this is the first time to study static VMPRSA problem in EODCNs. The major contributions of this work are: (i) The VMPRSA problem is formally stated and its NP-hard is proved. (ii) An integer linear programming (ILP) model for the static VMPRSA problem is presented, which can optimally assign the VMs for multi-tenant and the concomitant traffic demands between VMs. (iii) An efficient two-tier heuristic algorithm framework, the first tier is VM placement and the second is routing and spectrum assignment (VMP + RSA), is proposed to solve the VMPRSA problem. (iv) Two VM placement algorithms are proposed, namely random placement (RP) algorithm and residual node capacity priority (RNCP) algorithm. The simulation results indicate that the ILP model achieves the optimal solution, and between the two heuristic algorithms, the RNCP algorithm obtains the better sub-optimal solution.

2 Virtual Machine Placement and Routing Spectrum Assignment

In this section, the overview of traffic demands determined by placing tenant VMs on a set of physical servers (hereinafter referred to servers) is introduced.

2.1 Traffic Demands Driven by VM Placement

A slot is used to refer to one capacity to accommodate a VM placement on servers. A top of rack (ToR) switch has s servers and each server has multiple slots, and each slot can be occupied by any VMs. A scenario where there are m tenants and n slots is considered. For a valid VM placement for multi-tenant, the sum of VM demands (in terms of CPU/memory) of m tenants should be smaller than or equal to n. As shown in Fig. 1, there are 3 tenants, and the sum of VM demands of 3 tenants is 12. The network has 12 slots which can satisfy the VM demands of 3 tenants.

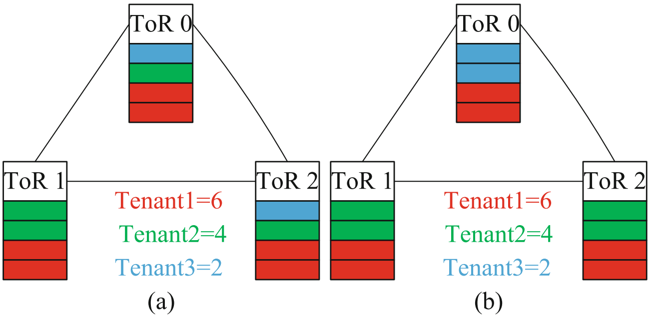

Fig. 1. Two schemes of VM placement.

A tenant demand needs some VMs, and each VM requires the network bandwidth to exchange data to all other VMs. VM placement determines the size of corresponding traffic demands and the switching method (i.e., electrical switching or optical switching) which is reflected into two aspects. On the one hand, if the VMs of a tenant are placed in a ToR switch, VMs send or receive data through electrical switching, which no cross-ToR traffic is produced. However, if the VMs of a tenant are placed into more ToR switches, cross-ToR traffic demands of all-to-all for the tenant would be produced. In this paper, cross-ToR traffic demands of all-to-all (hereinafter referred to traffic demands) using optical switching is focused. Note that the source node and the destination node of traffic demands are in respective of network nodes (i.e., ToR switch), not VMs themselves.

Then, how to quantify each traffic demand for a tenant is explained. When a tenant requires $m + n$ VMs, assume that m VMs are placed into a ToR switch and n VMs are placed into another ToR switch, therefore each ToR switch require m units of bandwidth and n units of bandwidth to send or receive data, respectively. If a ToR switch communicates with the other, the bandwidth will be limited to $\min(m, n)$ units. As shown in Fig. 1(a) which illustrates a scheme of VM placement, tenant 1 requires 6 VMs placed into 3 ToR switches, and each ToR switch has 2 VMs, therefore each ToR switch will require 2 units of bandwidth to exchange data. The communication bandwidth among ToR 0, ToR 1 and ToR 2 switches will be bounded to $\min(2, 2) = 2$ units; for tenant 2, the bandwidth between ToR 0 and ToR 1 is $\min(1, 2) = 1$ unit. In Fig. 1 (b), for tenant 3, all the VMs are placed in ToR 0, thus no traffic is yielded.

2.2 Virtual Machine Placement and Routing Spectrum Assignment

In this subsection, the static VMPRSA problem is formally stated in EODCNs. Define a network as $G(N, C, E, S)$, where N represents the set of nodes, C =

$\{C_0, C_1, ..., C_{|N|-1}\}$ is the set of capacity of nodes in N, E is the set of bidirectional fiber links between nodes in N, and S represents the set of spectrum slots on each fiber. The network node set has N nodes, it thus has $\sum_{i=0}^{|N|-1} C_i$ node capacity at all.

Definition. Static VMPRSA problem - given a network $G(N, C, E, S)$, and a predefined set of tenant demands (i.e., CPU/memory demands) with $|T|$ tenants ($T = \{T_0, T_1, ..., T_{|T|-1}\}$), where T_k represents that tenant k request T_k CPU/memory units, and satisfies $\sum_{k=0}^{|T|-1} T_k \le \sum_{i=0}^{|N|-1} C_i$. Integer variable $Allo_i^k$ represents that the size of VMs are placed in node i for tenant k. The traffic demand of any two nodes for tenant k is $\min(Allo_i^k, Allo_j^k)$. It is possible to establish each spectrum path in tenant k traffic demands and all tenants traffic demands using consecutive spectrum slots, and satisfy spectrum continuity constraint?

Theorem. The Static VMPRSA problem is NP-hard.

Proof. The typical RSA problem in elastic optical network is a special instance of the VMPRSA problem, in which traffic demands are fixed and no node capacity is considered [7]. Since the RSA problem alone is NP-hard, our claim holds.

3 ILP Model

In this section, an ILP mathematical model is formulated with the objective of minimizing the spectrum used to serve the multi-tenant demands, while the node capacity should satisfy the VM placement.

3.1 Notations and Variables

- T, the set of tenant demands, the element T_k represents the size of VM demands for tenant k;
- C_i, the capacity of node i, and each node has the same capacity for simplification;
- $Traf_{i,j}^k$, a decision integer variable that represents the traffic demand of tenant k between node i and j;
- $F_{i,j,m,n}^{w,k}$, a boolean variable, is equal to 1, if there is a spectrum path using spectrum slots w to satisfy the traffic demand of tenant k between node-pair (i, j) going from node m to node n, and 0 otherwise;
- M, an integer variable represents the maximum utilized spectrum slots for all the tenant traffic demands in the network. For a valid assignment, $|S|$ should always be bigger than or equal to M.

3.2 Objective and Constraints of the VMPRSA Problem

Since more spectrum slots used on a fiber signify more cost on the fiber and further need more corresponding switching equipment and power consumption, the objective of this model is to minimize the maximum spectrum slots utilized among all the fibers to serve all tenants traffic demands. Such traffic demands are determined by VM placement which should be satisfied by the node capacity. To simply our model, the guard-bandwidth between optical routing is not considered. Equation (1) is employed to represent object function.

$$Minimine\ M \tag{1}$$

subject to the following constraints:

$$Allo_i^k \leq T_k, \forall i, k \tag{2}$$

$$\sum_i Allo_i^k \leq T_k, \forall k \tag{3}$$

Equations (2) is VM placement constraints for tenants. It denotes that the size of VM placement for each tenant in each node can not exceed the tenant demand T_k, while the sum of such size in all nodes should equal T_k, which is guaranteed by (3).

$$\sum_k Allo_i^k \leq C_i, \forall i \tag{4}$$

Each node can be shared by multi-tenant, thus capacity provided by each node for multi-tenant VM placement does not exceed the node capacity, as shown in (4).

$$Traf_{i,j}^k = \min(Allo_i^k, Allo_j^k), \forall i, j, k \tag{5}$$

The traffic demand between any two nodes for each tenant is determined by the smaller size of tenant VM placement, as shown in (5).

$$F_{i,j,m,n}^{w,k} * w \leq \text{M}, \forall i, j, m, n, w, k \tag{6}$$

Cost function is shown in (6), which obtains the maximum utilized spectrum slots on each fiber.

$$\sum_{w,j=n,m} F_{i,j,m,n}^{w,k} = Traf_{i,j}^k, \forall i, j, k \tag{7}$$

$$\sum_{w,i=m,n} F_{i,j,m,n}^{w,k} = Traf_{i,j}^k, \forall i, j, k \tag{8}$$

$$\sum_{w,i=j} F_{i,j,m,n}^{w,k} = 0, \forall m, n, k \tag{9}$$

The traffic demand between node i and node j for tenant k should be exactly added at node i and dropped at node j, which are guaranteed by (7) and (8), respectively. Equation (9) makes sure that no traffic is required at the same node.

$$\sum_{i,j} F_{i,j,m,n}^{w,k} \leq 1, \forall m, n, k, w \tag{10}$$

One spectrum slot can only be used for satisfying one traffic demand for multi-tenant, which is specified by (10).

$$\sum_{w,j \neq n,m} F_{i,j,m,n}^{w,k} - \sum_{w,i \neq n,p} F_{i,j,n,p}^{w,k} = 0, \forall i, j, n, k \tag{11}$$

The spectrum continuity constraint guarantees that the spectrum path of a traffic demand should use the same spectrum through its routing path, which is shown in (11).

$$(F_{i,j,m,n}^{w,k} - F_{i,j,m,n}^{w,k+1}) * (-B) \geq \sum_{\overline{w} \in \{w+2, Cap\}} F_{i,j,m,n}^{\overline{w},k}, \forall i, j, m, n, w, k \tag{12}$$

$$(F_{i,j,m,n}^{w,k} - 1) * B + Traf_{i,j}^{k} \leq \sum_{\overline{w} \in \{1, Cap\}} F_{i,j,m,n}^{\overline{w},k}, \forall i, j, m, n, w, k \tag{13}$$

The spectrum consecutiveness constraint is shown in (12), which means that if $F_{i,j,m,n}^{w,k}=1$ and $F_{i,j,m,n}^{w+1,k}=0$, all the spectrum higher than $w+1$ will not be used for the spectrum path of node-pair (i, j) for tenant k on link $m-n$. Equation (13) guarantees that size of consecutive spectrum is $Traf_{i,j}^{k}$ if $F_{i,j,m,n}^{w,k}=1$. In above both equations, a large number B is introduced to realize the if-then relationship.

4 Heuristic Algorithm

The proposed ILP model is NP-hard and is tractable only when the problem size (e.g., the number of tenants, network topology) is small. For a large scale problem, heuristic algorithms are resorted to obtain sub-optimal solutions within reasonable time.

4.1 Two-Tier Algorithm Framework

Since the traffic demands of a tenant are driven by VM placement of the tenant, a two-tier heuristic algorithm framework is proposed to apply the problem.

The two-tier algorithm framework is consisted of VM placement and routing spectrum assignment (VMP + RSA), as shown in **Algorithm 1**. The first tier is VM placement of multi-tenant VM demands, which can produce the corresponding traffic demands and is shown on line 1, and the second tier is to execute RSA algorithm to server all the traffic demands, which is illustrated in the algorithm from line 2 to the end of it.

Since the objective of heuristic algorithms is dominated by the traffic demands of multi-tenant, which are determined by VM placement of multi-tenant, we mainly focus on the VM placement algorithms. The same RSA algorithm is employed to evaluate the network performance of different VM placement algorithms. The key idea of RSA algorithm is the bigger traffic demand with the higher priority, and each traffic demand choose the lowest starting spectrum from K paths which use the K-shortest path algorithm.

Algorithm 1 VMP+RSA algorithm framework

VMs placement algorithm
Input: $G(N,C,E,S), T_k, C_i, i \in N, TenTraf = \varnothing$
Output: $TenTraf = \{Traf_{i,j}^k\}, W = 0$
RSA algorithm
Input: $TenTraf = \{Traf_{i,j}^k\}, W = 0$
Output: W
Sort traffic demands of all tenants according to their sizes $Traf_{i,j}^k$ in descending order in a queue Q
while $Q \neq \varnothing$ **do**
Pick the top traffic demand of (i,j) pair in the Q, and search the lowest available starting spectrum m among all K paths;
Select the path with the lowest starting spectrum m within K paths, and assign the spectrum for it. If several paths have the same m, select the first;
$M = Traf_{i,j}^k + m$
If $W < M$ **then**
$W = M$
end if
Delete the top traffic demand of (i,j) pair from the queue Q, and update the network state;
end while

4.2 Random Placement

For the problem of VM placement, a naive approach is random placement (RP) algorithm, i.e., for a tenant VM demand, one node or more nodes are randomly chosen to accommodate all the VMs for the tenant. A tenant demand with more VMs has high priority, since which can avoid the tenant VMs allocated into different nodes to minimize the number of cross-ToR traffic demands. Furthermore, to achieve the goal, each node should accommodate VMs for a tenant with the maximum node residual capacity. Specially, If $T_k \geq C_i$, then $T_k = T_k - C_i$ and C_i=0. Else, $T_k = 0$ and $C_i = C_i - T_k$. From the algorithm, the next tenant VMs will not be allocated until the current tenant VMs have been allocated.

4.3 Residual Node Capacity Priority

However, the bigger traffic demands (in size) could be produced by the RP algorithm due to the randomness of VM placement. In EODCNs, the spectrum can not be evenly utilized by bigger traffic demands, which can make utilized spectrum bigger. Therefore, we should lower the size of traffic demands, which can result in the reduction of average size of traffic demands. As we known, when VMs of a tenant are allocated into more than one node, the traffic demands will appear, and their sizes are determined by the smaller size of VM placement between two nodes. Therefore, residual node capacity priority (RNCP) algorithm is proposed. The key idea of the algorithm is to place all the VMs of a tenant to a node as much as possible, and no traffic demand is yielded; when no node can not accommodate all the VMs for a tenant, a node with the largest residual capacity is employed to accommodate VMs as much as possible and pick a node with smallest residual capacity to accommodate the residual VMs.

5 Performance Evaluation

In this section, the optimal result of the proposed ILP model and sub-optimal results of the two heuristic algorithms are evaluated. The ILP model and the two heuristic algorithms are enforced by using the CPLEX 12.5 and Visual Studio 2015 C++ simulation platform, respectively. Without loss of generality, 5-node and fat-tree networks are employed as EODCNs topology to implement our simulations, as shown in Fig. 2. A ToR switch is represented by a node of network, and servers are omitted for simplification. Traffic demands are launched and terminated by network nodes.

Assume that the width of a spectrum slot is 12.5 GHZ and the guard bandwidth is not considered. For a valid placement, a network has $\sum_{i=0}^{|N|-1} C_i$ capacity and can accommodate $|T|$ tenants. Suppose that a tenant VM demand is a random integer, and $|T|$ random integers whose sum is equal to $\sum_{i=0}^{|N|-1} C_i$, can contribute to certain standard variance. Random integers with the same standard variance should be considered for fairness by changing $|T|$. However, it is very difficult to generate random integers with given standard variance. Therefore, predefine that each tenant has the same VM demand with the same standard variance (i.e., 0 standard variance) for simplification. Furthermore, since the traffic demands are symmetric, we only consider the unidirectional traffic demands.

5.1 ILP and Heuristic Algorithms Under a Small Topology

To compare the performance of the ILP model with the heuristic algorithms, they are enforced on a 5-node network with six bidirectional links as shown in Fig. 2(a). The metrics here include spectrum utilized to sever all the traffic demands, and average size of traffic demand.

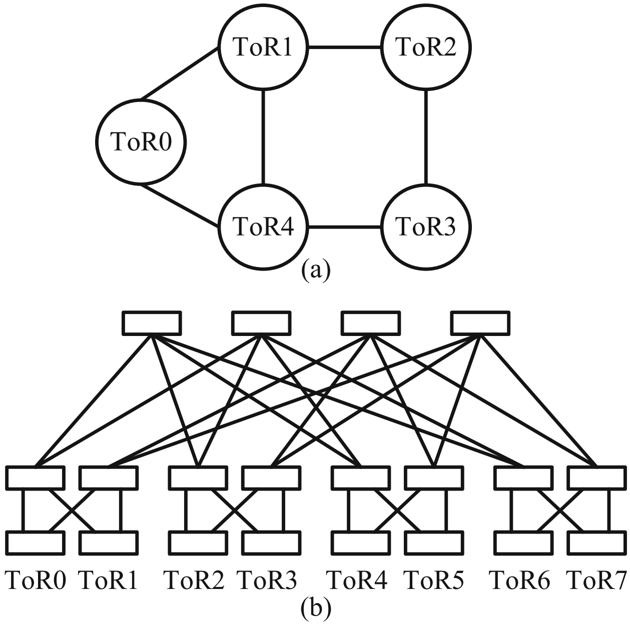

Fig. 2. (a) 5-node topology; (b) fat-tree topology.

The spectrum required to serve all the traffic demands driven by VM placement as $|T|$ increases is investigated and the results are shown in Fig. 3. In the network, each node has the same capacity with 12 slots, the network thus has a capacity of 60 slots. It is clearly that the spectrum required decreases along with the increasing $|T|$. This is because each tenant VM demand T_k decreases along with the increase in $|T|$, which can produce smaller traffic demands and thus lower the utilized spectrum. The ILP model gets the best solution compared with the two heuristic algorithms. RNCP is better than RP. This is because RNCP tends to consider the node with residual capacity, leading to a smaller size of these traffic demand. Please note that no spectrum is required when $|T|$=5 with T_k=12 for each tenant, since each node can exactly accommodate a tenant VMs, which produces no traffic demand for any tenants. It can be clearly seen that when $|T|$= 2, ILP outperforms RNCP and RP by up to 42.8% and 71.4%, respectively.

The average size of traffic demands driven by VM placement is studied by changing $|T|$ and the results are represented by Fig. 4. The main trend is decreasing due to the degradation of T_k. It is obvious that the ILP model gets the smallest value. RNCP is better than RP, because the former takes the residual node capacity into account when carry out multi-placement for a tenant VMs, while the RP algorithm considers the random node. When $|T|$= 2, ILP is better than RNCP and RP by up to 40% and 60%, respectively.

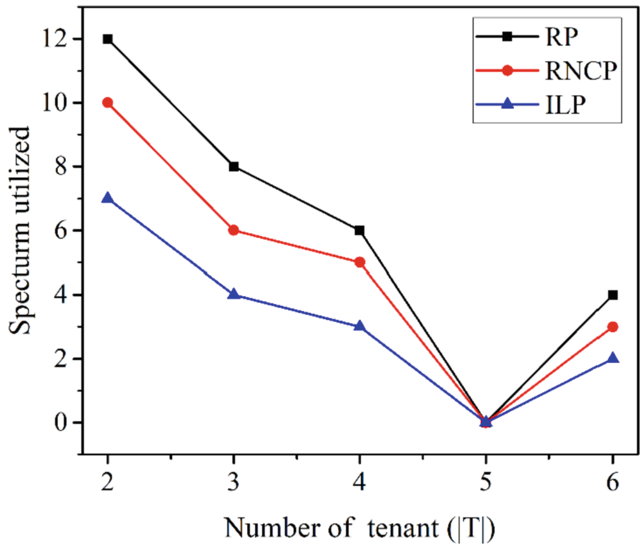

Fig. 3. Spectrum utilized versus number of tenant in a 5-node topology.

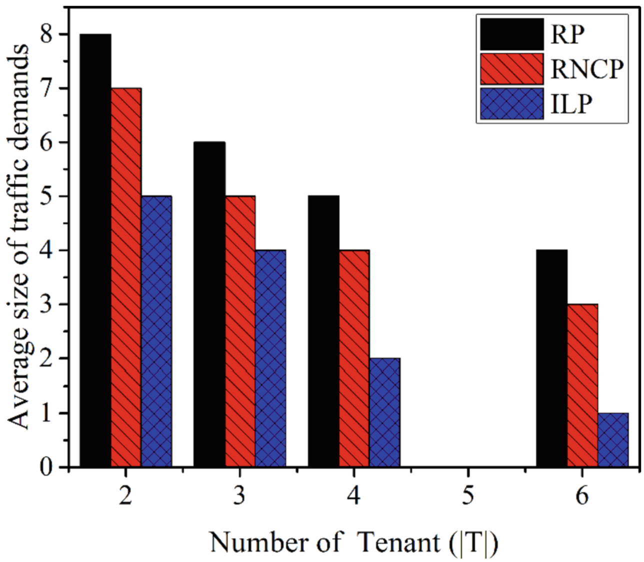

Fig. 4. Average size of traffic demands versus bandwidth granularity in a 5-node topology.

5.2 Heuristic Algorithms in a Large Topology

The fat-tree network with 8 nodes is used as EODCNs topology shown in Fig. 2(b). In the network, each node has the same capacity with 140 slots, the network thus has a capacity of 1120 slots.

The spectrum utilized to server all the traffic demands driven by VM placement as $|T|$ increases is exhibited by Fig. 5. The main tread is declining, since more tenants VMs can be fully placed into a node. The RNCP algorithm obtains the better sub-optimal solution, because it considers the residual node capacity when carry out multi-placement for a tenant VMs, leading to the small traffic demands. When the number of tenant $|T|$ is 60, RNCP outperforms RP by up to 38.4%.

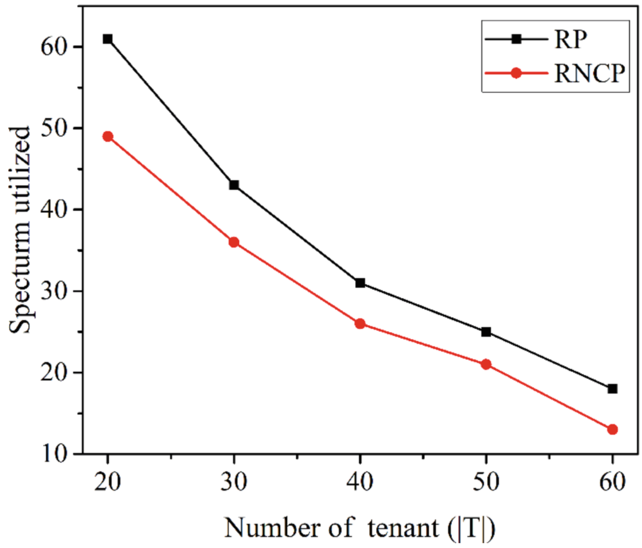

Fig. 5. Spectrum utilized versus number of tenant in a fat-tree topology.

The average size of traffic demands is also considered by different $|T|$ and the results are depicted in Fig. 6. The main trend is decreasing, since the number of tenants increase and the T_k decreases. It is obvious that RNCP is better than RP. When $|T| = 60$, RNCP outperforms RP by up to 60%.

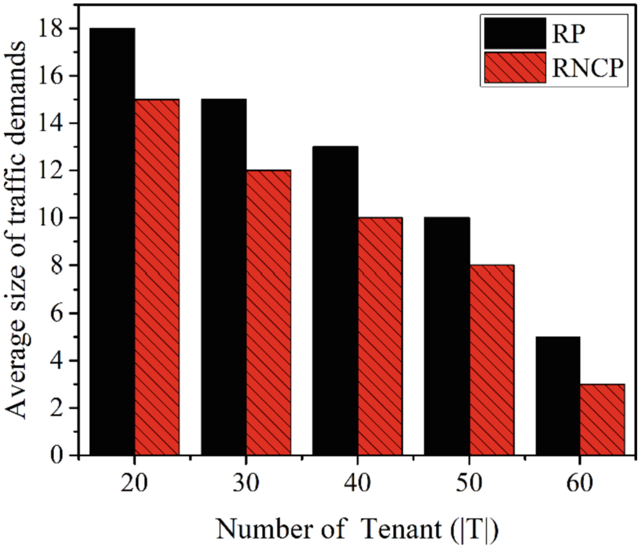

Fig. 6. Average size of traffic demands versus bandwidth granularity in a fat-tree topology.

6 Conclusion

In this paper, the static virtual machine placement and routing spectrum assignment (VMPRSA) problem is investigated in EODCNs. The target of the problem is to minimize the utilized spectrum to serve all the traffic demands driven by VM placement of multi-tenant, while the network should have enough node capacity to satisfy the VM placement. The VMPRSA problem is formally stated and its NP-hard is proved. An ILP model is presented to solve the problem. Since the ILP model can not scale to a big network, a two-tier heuristic algorithm framework is proposed to apply the problem, i.e., the first tier is VM placement and the second tier is routing and spectrum assignment (VMP + RSA). Two VM placement algorithms are proposed, the RP algorithm and the RNCP algorithm. The simulation results show that the ILP model provides the optimal solution, and the RNCP algorithm achieves the better sub-optimal solution. The dynamic VMPRSA problem will be studied in future work.

References

1. Al-Fares, M., et al.: A scalable, commodity data center network architecture. ACM SIGCOMM Comput. Commun. Rev. **38**(4), 63–74 (2008)
2. Singla, A., et al.: Proteus: a topology malleable data center network. In: The 9th USENIX Symposium on Networked Systems Design and Implementation, pp. 225–238 (2010)
3. Farrington, N., et al.: Helios: a hybrid electrical/optical switch architecture for modular data centers. ACM SIGCOMM Comput. Commun. Rev. **41**(4), 339–350 (2011)
4. Wang, G., et al.: C-Through: part-time optics in data centers. ACM SIGCOMM Comput. Commun. Rev. **41**(4), 327–338 (2011)

5. Porter, G., et al.: Integrating microsecond circuit switching into the data center. ACM SIGCOMM Comput. Commun. Rev. **43**(3), 447–458 (2013)
6. Chen, K., et al.: OSA: an optical switching architecture for data center networks with unprecedented flexibility. IEEE/ACM Trans. Networking **22**(2), 498–511 (2014)
7. Jinno, M., et al.: Spectrum-efficient and scalable elastic optical path network: architecture, benefits, and enabling technologies. IEEE Commun. Mag. **47**(11), 66–73 (2009)
8. Liu, A., Sun, Y., Ji, Y.: Bandwidth reservation for tenants in reconfigurable optical OFDM datacenter networks. IEEE Photonics J. **10**(5), 7906616 (2018)
9. Ballani, H., et al.: Towards predictable datacenter networks. ACM SIGCOMM Comput. Commun. Rev. **41**(4), 242–253 (2011)