Efficient Algorithm for Energy-Aware Virtual Network Embedding

Shuxian Jia
the School of Computer Software Technology, Zhejiang University, Ningbo 315048, China.

Guiyuan Jiang
the School of Computer Science and Engineering, Nanyang Technological University, 639798, Singapore.

Peilan He
the School of Computer Science and Engineering, Nanyang Technological University, 639798, Singapore.

Jigang Wu
the State Key Laboratory of Computer Architecture, Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China.

Follow this and additional works at: https://tsinghuauniversitypress.researchcommons.org/tsinghua-science-and-technology

Part of the Computer Sciences Commons, and the Electrical and Computer Engineering Commons

Recommended Citation
Efficient Algorithm for Energy-Aware Virtual Network Embedding

Shuxian Jia, Guiyuan Jiang, Peilan He, and Jigang Wu*

Abstract: Network virtualization is a promising approach for resource management that allows customized Virtual Networks (VNs) to be multiplexed on a shared physical infrastructure. A key function that network virtualization can provide is Virtual Network Embedding (VNE), which maps virtual networks requested by users to a shared substrate network maintained by an Internet service provider. Existing research has worked on this, but has primarily focused on maximizing the revenue of the Internet service provider. In this paper, we consider energy-aware virtual network embedding, which aims at minimizing the energy consumption for embedding virtual networks in a substrate network. In our optimization model, we consider energy consumption of both links and nodes. We propose an efficient heuristic to assign virtual nodes to appropriate substrate nodes based on priority, where existing activated nodes have higher priority for hosting newly arrived virtual nodes. In addition, our proposed algorithm can take advantage of activated links for embedding virtual links so as to minimize total energy consumption. The simulation results show that, for all the cases considered, our algorithm can improve upon previous work by an average of 12.6% on acceptance rate, while the consumed energy can be reduced by 12.34% on average.

Key words: virtualization technology; virtual network embedding; energy efficient; optimization algorithm

1 Introduction

With the development of data centers and cloud computing, network service requests from end users impose various requirements on Quality-of-Service (QoS) guarantees. The ever-increasing requirements placed on the network are fueling its evolution to architectures that make more efficient use of available network resources, and promote service innovations. Internet Service Providers (ISPs) have to satisfy individual needs for their customers, and hence are impelled to use different protocol stacks and provide custom services and network resources.

In recent years, the networking research community has become increasingly interested in efficient and flexible management approaches for resource allocation. Virtualization technology, which utilizes critical resources by creating virtual (rather than actual) versions of the corresponding resources, such as operating systems, servers, storage devices, and network resources, provides new opportunities for improving resource efficiency. In particular, the concept of network virtualization[1–3], allowing customized Virtual Networks (VNs) to be set up on a shared physical infrastructure, has been proposed as a promising approach for the future Internet, which aims to overcome the resistance of the current Internet to architectural change.

Virtualization is an abstract concept that hides the complexity of underlying substrate network details, thereby allowing ISPs to cope with the wide range of customer requirements. In network virtualization, Infrastructure Providers (InPs) divide their resources

* To whom correspondence should be addressed.
Manuscript received: 2016-02-21; accepted: 2016-03-22
into chunks, called VNs, which are allocated to SPs. Thanks to virtualization, the resource chunks are isolated from each other, so the service networks behave as if they were independent, though they share the same substrate infrastructure. Large enterprises like Google, Facebook, and Amazon actually utilize virtualized data centers for large-scale computation and storage\footnote{4}.

Due to the advantages of the technique, it has received significant attention, since it makes it possible to set up a cost-effective data center infrastructure for storing large volumes of data and to host large-scale service applications. However, this technology requires efficient algorithms, commonly called “Virtual Network Embedding (VNE)’ algorithms, to instantiate virtualized networks on a substrate infrastructure, and optimize the layout for service-relevant metrics.

Previous work has shown that the VNE problem is NP-hard, as it can be reduced to a multi-way separator problem\textsuperscript{15–7}. The VNE problem has received extensive investigation from different perspectives such as load balancing in the substrate network\textsuperscript{8} and efficient use of resources\textsuperscript{9}. The challenges are principally caused by the high computational complexity\textsuperscript{10,11}. Most proposals have focused on a two-phase embedding approach, first by preselecting the mapping of virtual nodes, and second by assigning virtual links to substrate paths. This approach ignores the benefit of joint design of node mapping and link mapping through efficiency coordination. On the other hand, many proposals in the literature have therefore focused on restricting the problem to particular scenarios, such as assuming infinite network capacities in order to gain obvious admission control, ignoring either virtual link or node QoS requirements\textsuperscript{12}, relaxing embedding path-splitting constraints\textsuperscript{13}, focusing on specific VN topologies\textsuperscript{12}, and designing greedy embedding algorithms\textsuperscript{12,14}.

Network virtualization techniques can also lead to significant energy savings by means of resource consolidation\textsuperscript{15}. Existing Energy-Efficient VNE (EEVNE) algorithms use different approaches to minimize energy consumption. Razzaq et al.\textsuperscript{16} proposed a Baseline Approach (BLA) to deal with assigning virtual nodes to any substrate nodes that satisfy the demand. A path-based model called P-VNE was presented by Hu et al.\textsuperscript{17} for the VNE problem. Xu et al.\textsuperscript{18} proposed an efficient embedding algorithm, which explores periodic resource demands of virtual networks and employs a new embedding metric. Guan et al.\textsuperscript{19} built a novel model of a VNE problem to minimize usage in data centers. Oliveira et al.\textsuperscript{20} presented an Opportunistic Resilience Embedding (ORE), which protects virtual links against substrate network disruptions.

Most existing EEVNE algorithms try to minimize energy consumption by resource consolidation. However, due to the relationship between the energy consumption of nodes and CPU utilization, virtual node embedding has great impact on the energy consumption of substrate networks, which is not considered in existing EEVNE algorithms. In this paper, we investigate the problem of energy-efficient virtual network embedding and propose an efficient algorithm to obtain minimization of energy consumption of virtual node embedding under the constraints of virtual link embedding. Specifically, we formulate virtual network embedding as an optimization problem for minimizing energy consumption and propose an efficient heuristic algorithm to reduce energy consumption for the substrate network. We conduct extensive experiments to evaluate the performance of our proposed algorithm.

The rest of this paper is organized as follows. We introduce related definitions and important notations, and formulate the EEVNE problem in Section 2. The proposed algorithm is introduced in Section 3, and Section 4 presents the experimental results with analysis. We conclude our work in Section 5.

2 Preliminaries

In this section, we present the definitions and important notations, then formulate virtual network embedding as an optimization problem for minimizing energy consumption.

2.1 Notation

A substrate network is modeled as an undirected weighted graph $G_P = (N_P, L_P)$, where $N_P$ is the set of physical nodes and $L_P$ is the set of physical links. Similar to the substrate network, a virtual network is modeled as an undirected graph, which is denoted as $G_V = (N_V, L_V)$, where $N_V$ is the set of virtual nodes and $L_V$ is the set of virtual links. Let $m_{SN}$ be the number of physical nodes, i.e., $m_{SN} = |N_P|$, and $n_{VN}$ is the number of virtual nodes, i.e., $n_{VN} = |L_P|$. $|L_V|$ is the total number of the set of virtual links.

Each physical node $n_p \in N_P$ is associated with $\text{CP}(n_p)$ units of CPU resources, and $\text{ACP}(n_p)$ units are currently available for accepting new computation
Let $C_b$ denote the base energy consumption, $C_m$ be the total energy consumption, and $C_{\text{curr}}$ represents the total energy consumption for all the nodes that have been allocated into the physical networks. Let $\text{ABR}(n_s)$ be the amount of unoccupied bandwidth of all the outgoing links of substrate node $n_s$. Each substrate link $(i, j)((i, j) \in L_p)$ from physical node $i$ to $j$ is associated with the amount of energy consumption $C_n(l_p)$.

For every virtual node $n_v$, the demanding CPU resource is denoted as $\text{DCP}(n_v)$, while the required bandwidth for virtual link $l_v$ is denoted as $\text{ABR}(l_v)$. The amount of required bandwidth of all outgoing links of a given virtual node $n_v$ is denoted as $\text{DBR}(n_v)$.

### 2.2 Problem description

In this paper, we focus on the impact of virtual node mapping on the energy consumption of a substrate network. We keep the same model and assumption as previous work\cite{21}. The physical network is quantified as follows.

Let $\eta_i$ be the rate of CPU utilization of a physical node. Since the CPU utilization is a main contributor to the energy consumption of a server, the $i$-th node energy consumption $E_i$ can be calculated as $C_b + C_1 \cdot \eta_i$, where $C_b$ is the base energy consumption of the substrate node, $C_m$ is the total energy consumption of a substrate node.

Let the binary variable $a(e)$ be 1 if the node or link $e$ is activated, and $a(e) = 0$ otherwise. Thus, if a node or link is not activated, its energy consumption is 0. The energy consumption model of the nodes and links of a physical network is calculated as follows.

$$E_{SN} = \sum_{n_v \in N_v} a(n_v) \cdot (C_b - (C_m - C_b) \times \eta_{n_v}) + \sum_{l_p \in L_p} a(l_p) \cdot C_n(l_p)$$

(1)

In order to reduce the energy consumption, we should consider the energy consumption of both nodes and links. On the other hand, if the available CPU resource $\text{DCP}(n_v)$ is less than the required CPU resource $\text{DCP}(n_v)$, or the available bandwidth $\text{ABR}(n_s)$ is less than required $\text{DBR}(n_s)$, then the mapping is not successful. Otherwise we denote the energy cost per CPU of mapping $n_1$ into $n_2$ as follows. Let $\text{cost}_{ij}$ be the energy cost per CPU of embedding $i$ into $j$. Note that if a physical node is not active, we need to set the $\text{cost}_{ij}$ to a large value. Hence, $\text{cost}_{ij}$ is calculated as

$$\text{cost}_{ij} = \begin{cases} 
\frac{1}{\text{CP}(n_p)}, & \text{if physical node } n_p \text{ is active;} \\
\frac{1}{\min \{\text{CP}(n_p)\} - 0.0001}, & \text{otherwise}
\end{cases}$$

(2)

According to the aforementioned theory, the key to reducing the energy cost per CPU of mapping $i$ to $j$ is allocating priority in the VEN problem model. In this section, we propose a novel way to calculate the priority for selecting a physical node.

$$W_{\text{cpu}}(n_p) = \alpha \times \text{CP}(n_p) + \beta \times \text{DCP}(n_p) + (1 - a(n_p)) \times \text{MAX}$$

(3)

where $\alpha$ and $\beta$ are weight factors, and $\text{MAX}$ is a constant integer. Here, we set $\text{MAX}$ to be the biggest capacity of all substrate nodes. It is obvious that increasing the parameter $\alpha$ can improve the efficiency for embedding a computation-intensive virtual network, while increasing the parameter $\beta$ is more useful for embedding communication-intensive virtual networks.

Through the above analysis, the EEVEN problem can be formulated as follows.

Objective:

$$\sum_{n_v \in N_v} a(n_v) \cdot (C_b - (C_m - C_b) \times \eta_{n_v}) + \sum_{l_p \in L_p} a(l_p) \cdot C_n(l_p)$$

(4)

s.t.

$$\sum_{n_v \in N_v} \text{DCP}(n_v) \cdot x(n_v, n_p) \leq \text{CP}(n_p)$$

(5)

$$\sum_{l_v \in L_v} \text{DBR}(l_v) \cdot x(n_v, n_p) \leq \text{ABR}(n_p)$$

(6)

$$\sum_{i \in N_v} x(i, u) \leq 1$$

(7)

$$\sum_{u \in N_p} x(i, u) = 1$$

(8)

where the binary variant $x(i, u)$ is defined as follows.

$$x(i, u) = \begin{cases} 
1, & \text{if node } i \text{ is embedded to node } u; \\
0, & \text{otherwise}
\end{cases}$$

(9)

Constraint (5) guarantees that the total required CPU capacity of virtual nodes mapped to physical nodes always be equal to or less than the total required CPU capacity of virtual nodes mapped to physical nodes, which must always be equal to or less than the available computational resources on it. Similarly, another constraint (6) ensures that all of the bandwidth...
demanded by the virtual links will not exceed the unoccupied bandwidth of all the outgoing links of a substrate node. Constraint (7) is used to ensure that every i will be linked to a substrate node u. Constraint (8) is employed to ensure that each substrate node allows no more than one virtual node.

3 Proposed Algorithms

In this section, we proposed an efficient heuristic algorithm shown in Algorithm 1, called EH_Alg, to minimize the energy consumption of virtual node embedding under the constraints of virtual link embedding.

In our approach, EH_Alg determines how to assign a virtual node according to the following principles.

(1) It uses $W_{cost}(n_p)$ to measure the priority of a substrate node of the physical network, to accommodate target virtual nodes, according to the total CPU resources and unoccupied CPU capacities. In particular, the proposed $W_{cost}(n_p)$ allows embedding more than one virtual node into the same physical node.

(2) After finding a substrate node $j$, we consider that a virtual node $i$ is mapped into $j$. If the substrate node $j$ still has the unoccupied CPU resources, then we try to embed the neighbor node of node $i$ into substrate node $j$. If the virtual node $i$ has more than one neighbor node, the algorithm chooses a node according to the amount of required bandwidth between $i$ and its neighbors. The node with large bandwidth requirements takes higher priority over those with low bandwidth needs. The process repeats until substrate node $j$ is overloaded. This process minimizes the use of communication resources.

The detailed algorithm is shown in algorithm EH_Alg. For simplicity, we assume that all virtual nodes arriving within a particular timeframe are assigned to the substrate network, otherwise the virtual network will be rejected. The algorithm runs are simulated over 500 time slots, with 10 virtual nodes arriving in each time slot.

We employ an example to show the substrate network, the virtual network, and the process of embedding. Figure 1a illustrates a substrate network, where the integer in each rectangle represents the available CPU resources that the physical nodes can offer, and the value of each link indicates the available bandwidth. Figure 1b shows a virtual network consisting of 4 virtual nodes $u$, $v$, $w$, and $p$.

Algorithm 1 EH_Alg

Input: A set of virtual nodes at a each time slot and the physical network $G_P = (N_P, L_P)$;
Output: Map as many virtual nodes as possible into substrate network $G_P$ while simultaneously minimizing the total energy cost;
1: Begin
2: for each time slot do
3: for each unmapped virtual node $i$ do
4: compute $W_{cost}(n_p)$ for each substrate node $n_p$;
5: $j \leftarrow \arg \min_{n_p \in V_P} \{ W_{cost}(n_p) \}$;
6: Map virtual node $i$ to substrate node $j$;
7: while node $j$ can accommodate more nodes do
8: MS $\leftarrow$ the set of nodes mapped to node $j$, $NS \leftarrow$ unmapped neighbors of nodes in MS;
9: $c \leftarrow$ the virtual node in NS demand maximal bandwidth;
10: Map this virtual node $c$ to substrate node $j$;
11: end while
12: end for
13: end for
14: End

Fig. 1 Example of virtual network embedding.
p. Each virtual node requires a certain amount of CPU resources, denoted by the value inside the nodes symbol. The value of the link between two virtual nodes indicates the amount of required communications.

Figure 1c illustrates how to allocate the virtual network into the substrate network using the proposed approach. In the first step, we select physical node A in the substrate network according to the calculated priority. Then, the algorithm assigns node u to the substrate node A. At this stage, node A has enough CPU resources to host additional virtual nodes, whether to assign one of u’s neighbors to node A. Node v is mapped to substrate node A because there are many communication requests between node u and node v, so mapping them to the same substrate node can significantly reduce energy consumption. After virtual nodes u and v have been mapped to substrate node A, node A still has enough CPU resources to host other virtual nodes, as shown in Fig. 1c. The algorithm repeats this process and tries to map other nodes into substrate node A. If node A cannot host any more nodes, the algorithm will try to find another substrate node until all the virtual nodes are successfully embedded or the virtual network is rejected.

We now analyze complexity of algorithm EH_Alg. Since the algorithm EH_Alg is an online algorithm, we only analyze the time complexity of embedding a given virtual network $G' = (V', E')$, which contains $|V|$ virtual nodes and $|E|$ virtual links, into a given substrate network $G(V, E)$ consisting of $|V|$ physical nodes and $|E|$ physical links. Clearly, procedure MAP $(i, j)$, as shown in Algorithm 2, runs in at most $O(|V'| \times (|E| + |V|))$ time. This is because, when mapping node i into node j, there are at most $|V'|$ edges needed to be mapped, as node i has at most $|V'| - 1$ neighbors, and mapping one edge takes at most $O((|E| + |V|))$ time. For each virtual network, there are $O(|V'|^2 \times (|E| + |V|))$ nodes in total, and thus the time complexity is $O(|V'|^2 \times (|E| + |V|))$.

4 Simulation

In this section, we evaluate the proposed approach by comparing it to existing algorithms. Same as the previous work, we use the GT-ITM tool to generate the substrate network topology consisting of 100 physical nodes and 570 physical links, which corresponds to a medium-sized ISP. For each substrate node, the CPU resources are set following a uniform distribution from 50 to 100. Similarly, for each substrate link, the bandwidth also follows this range from 50 to 100. Both algorithms take the same input substrate network and virtual networks, and both repeat over 500 time slots.

In order to evaluate our proposed algorithm, we compare the EH_Alg with algorithm TR-CL from Ref. [21]. Referring to the previous research, we set the values of $C_b$, $C_m$, and $CP(n_p)$ to be 150, 300, and 15, respectively. When collecting simulation results, both $\alpha$ and $\beta$ are set to 0.5.

To compare the existing algorithms, we consider a certain status, named a “closely-saturated” state, which means the CPU resources of physical nodes are close enough to accommodate all the virtual nodes at a certain time slot. To evaluate this state, the CPU resources of the virtual nodes and the bandwidth of the virtual links are randomly distributed between 0 and 12. The evaluation results are shown in Figs. 2–6.

1. The revenue is defined as $DCP(n_s) + DBR(\lambda)$, which can be seen in Fig. 2. Our algorithm has great performance in terms of long-term average energy consumption, shown in Fig. 3.

2. Our algorithm EH_Alg significantly exceeds the TR-CL algorithm in terms of its acceptance ratio. In Fig. 4, the acceptance ratios of EH_Alg and TR-CL are 100% and 89%, respectively. Specifically, when a certain substrate node is still superior to the other substrate nodes after mapping virtual node

```
Algorithm 2 MAP(i, j)
Input: Virtual node i and the physical network $G_P = (N_P, L_P)$;
Output: Try to map each of node i to the substrate node j;
1: Begin
2: if j can host node i then
3: map i into j, and update node j;
4: else
5: reject the virtual network and update substrate network then break;
6: end if
7: while there is an unmapped link (i, k) for each mapped virtual node k do
8: find a shortest path between host(i) and host(j);
9: if there is no feasible path then
10: reject the virtual network and update substrate network then break;
11: else
12: map the link (i, k);
13: update information for all links on the path;
14: end if
15: end while
16: End
```
i.e., link power. It's easy to see that the energy consumption of both algorithms increases with the value of $C_n(l_p)$. The total energy consumption contains both server (node) computation and data communication on certain links. As previously discussed, the energy consumption of data communication on links can be reduced by mapping closely related virtual nodes (i.e., nodes that have considerable communication) to the same substrate node. Keeping this in mind, our proposed approach considers the communication demand among virtual nodes when performing virtual network embedding, which clearly outperforms the existing algorithm in terms of energy consumption.

(4) Our principle is that multiple virtual nodes can be mapped into a certain substrate node when the load of that physical node is less than a certain threshold $a$. $a$ is set to 70%, 80%, or 90%, as shown in Fig. 6. For example, when $a$ is set to 90%, the substrate network can hold all virtual nodes.

5 Conclusion

Virtual network embedding is a crucial problem in network virtualization. In this paper, the energy efficiency of virtual network embedding while
simultaneously minimizing total energy consumption is investigated. We model the VNE optimization problem as an Integer Linear Program that considers the impact of energy consumption on both nodes and links of a substrate network. We also propose a new heuristic algorithm to produce an efficient embedding solution based on the defined priority of substrate nodes and virtual nodes. Our experimental results verify the efficiency of the proposed algorithm.

Acknowledgment

This work was supported by the Specialized Research Fund for the Doctoral Program of Higher Education of China (No. 20131201110002), and the Key Laboratory of Computer Architecture Opening Topic Fund Subsidization (No. CARCH201303).

References

Guiyuan Jiang received the BS degree from Northwest University for Nationalities, China in 2008, the MEng degree from Tianjin Polytechnic University, China in 2011, and the doctoral degree from Tianjin University (TJU) in 2015, all in computer science and technology. He was an intern student at Nanyang Technological University (NTU) and National University of Singapore (NUS), respectively. Currently, he is working as a research fellow at School of Computer Science and Technology of NTU. His research interests include datacenter and cloud computing, fault-tolerant reconfiguration, hardware/software co-design, wireless sensor networks, and collaborative sensing for real-time traffic management.

Peilan He received the bachelor degree from Hainan University in 2012 and MEng degree in computer science and technology from Tianjin University, China in 2015. She is currently working towards the PhD degree at the School of Computer Science and Engineering, Nanyang Technological University, Singapore. Her research interests include intelligent transportation system, collaborative sensing for real-time traffic management, and data mining techniques for wireless sensor networks.

Shuxian Jia received the bachelor degree from Tianjin Polytechnic University, China, in 2015. Currently she is working towards the master degree at the School of Computer Software Technology, Zhejiang University (ZJU), China. Her research interests include datacenter and cloud computing, combinatorial optimization problems (designing efficient algorithm), and financial data analysis.

Jigang Wu received the BSc degree from Lanzhou University, China in 1983, and doctoral degree from University of Science and Technology of China (USTC) in 2000. He was with the Center for High Performance Embedded Systems, School of Computer Engineering, Nanyang Technological University, Singapore from 2000 to 2009, as a research fellow. He was with the School of Computer Science and Software Engineering as Dean and Tianjin distinguished professor, Tianjin Polytechnic University, China, from 2009 to 2014. Currently, he is with the School of Computer Science and Technology, Guangdong University of Technology. His research interests include reconfigurable VLSI design, hardware/software co-design, and combinatorial search.