

FLEXIBLE VIRTUAL ROUTING FUNCTION DEPLOYMENT IN NFV-BASED NETWORK WITH MINIMUM TOTAL NETWORK COST

Shin-ichiKuribayashi

Department of Computer and Information Science, Seikei University, Japan

ABSTRACT

In a conventional network, most network devices, such as routers, are dedicated devices that do not have much variation in capacity. In recent years, a new concept of Network Functions Virtualisation (NFV) has come into use. The intention is to implement a variety of network functions with software on general-purpose servers and this allows the network operator to select any capabilities and locations of network functions without any physical constraints.

This paper focuses on the deployment of NFV-based routing functions which are one of critical virtual network functions, and present the algorithm of virtual routing function allocation that minimize the total network cost. In addition, this paper presents the useful allocation policy of virtual routing functions, based on an evaluation with a ladder-shaped network model. This policy takes the ratio of the cost of a routing function to that of a circuit and traffic distribution in the network into consideration. Furthermore, this paper shows that there are cases where the use of NFV-based routing functions makes it possible to reduce the total network cost dramatically, in comparison to a conventional network, in which it is not economically viable to distribute small-capacity routing functions.

KEYWORDS

Virtual routing function, NFV, resource allocation, minimum total network cost.

1. INTRODUCTION

In recent years, a new concept of Network Functions Virtualization (NFV) [1]-[3] has been introduced. The main idea of NFV is the decoupling of physical network equipment from the functions that run on them. By virtualizing and consolidating network functions traditionally implemented in dedicated hardware and using cloud technologies, network operators expect to achieve greater agility and accelerate new service deployments while driving down both operational and capital costs.

The NFV implements a variety of virtual network functions with software on general-purpose servers, enabling the network operator to select any capabilities and locations of

network functions without any physical constraints. In addition, since NFV makes it easy to change functions and their locations and capacities dynamically to suit the usage condition, an economical network can be constructed flexibly. Therefore, the algorithm to determine the place where network functions are located and how much capacities of network functions flexibly are required is essential for economical network design.

The rest of this paper is organized as follows. Section 2 explains related works. Section 3 presents points of NFV-based routing function allocation. Section 4 proposes the algorithm of virtual routing function allocation with minimum total network cost. Section 5 evaluates the proposed algorithm using a ladder-shaped network model, which simulates the shape of Japan, and identifies the useful function allocation policy. Section 6 presents the conclusions. This paper is an extension of the study presented in [15]. Especially, this paper tries to minimize the total network cost, instead of the total power consumption.

2. RELATED WORK

Reference [2] summarizes the main differences between the conventional networks and NFV-based networks as follows:

- Decoupling software from hardware
- Flexible network function deployment
- Dynamic operation

The above features of NFV enable the network operator to select any capacities and locations of network functions without any physical constraints. Therefore, the efficient algorithm to determine the place where network functions are located and how much capacities of network functions are required is essential for economical network design.

The issue of placing network functions would be related to virtual network embedding (VNE) [4], [5], which is used to allocate physical resources for virtual nodes/links to construct virtual network. Through dynamic mapping of virtual resources onto physical hardware, the benefit gained from existing hardware can be maximized. Reference [6] proposes a virtual network function (VNF) chaining placement that combines location-routing and VNE problems, solving first the placement and then the chaining. Reference [7] focuses on a hybrid scenario where services are provided by dedicated physical hardware and virtualized service instances, and formulates the placement problem as an Integer Linear Program (ILP) with an objective of allocating a service chain onto the physical network minimizing the number of servers used. Reference [8] investigates the influence of virtualizing the S-GW and P-GW functions on the network load and network delay. It showed differences in performance when the functions were either fully virtualized and when their data and control planes were separated. The authors showed that four data centers are required to support a full virtualized gateways deployment for an exemplary gateway deployment in a US network. Reference [9] formulates the vDPI (virtual Deep Packet Inspection) placement as a cost minimizing problem, capturing the

different constraints the operator is facing, and solves it as an integer linear program (ILP).

References [10]-[12] propose the joint multiple resource allocation method in a cloud computing environment that consists of multiple data centers with processing ability and bandwidth to access them, and demonstrate by simulation evaluations that the proposed methods can reduce the total resource much, compared with the conventional allocation method. The proposed resource allocation model is related to the online bin packing problem with composite bins, in which two bins (processing ability bin and bandwidth bin) are consolidated.

As for dynamic allocation, Reference [13] evaluates the online virtual function mapping and scheduling problem and proposes a set of greedy algorithms. Reference [14] proposes the automatic placement of the virtual nodes and the allocation of network services on them, supported by a monitoring system that collects and reports on the behavior of the resources.

Most of these studies mainly try to embed the virtual network functions to physical resources economically and would be useful for the short-term design of the network, to use the existing resources efficiently. It is also required to expand the network or to add new network functions in the future. As for long-term design of the network, this paper tries to determine the optimal capacity and location of virtual network functions without any physical constraints. This should be necessary for the network operators to design their network economically. This paper also tries to show that the use of NFV-based routing functions makes it possible to reduce the total network cost dramatically, in comparison to a conventional network.

3. POINTS OF NFV-BASED ROUTING FUNCTION ALLOCATION

In general, routing functions would be allocated either in a centralized approach or in a distributed approach, as illustrated in Figure 1. There are three areas (#A, #B, #C) and a circle, \circ , indicates virtual routing functions deployed at each area in Figure 1.

If a distributed approach is adopted, virtual routing functions are deployed in each area and intra-area traffic does not need to go through routing functions outside the area concerned. This results in less cost of a circuit than in a centralized approach. In addition, packet transfer delay is shorter, thereby enhancing the quality of service. On the other hand, virtual routing functions are deployed only in Area #A in a centralized approach, resulting in less cost of virtual routing functions than in a distributed approach. In a conventional network, the centralized approach would be an economically viable option because capacity variations of routers are limited. Particularly, small capacity routers are not available. In NFV-based network, both centralized and distributed approaches could be considered as routing functions of any capacities can be allocated at any location economically. It is necessary to take these factors into consideration in considering the routing function allocation in an NFV-based network.

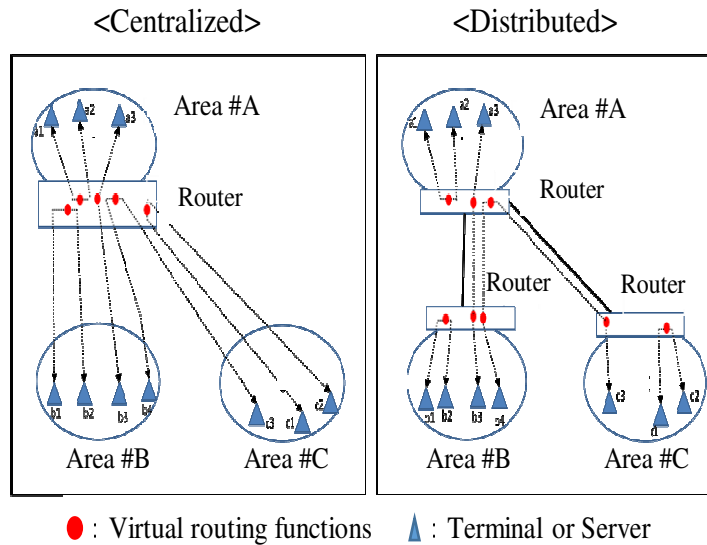


Figure 1. Two main approaches for virtual routing function allocation

4. ALGORITHM OF VIRTUAL ROUTING FUNCTION ALLOCATION WITH MINIMUM TOTAL NETWORK COST

4.1. TWO POSSIBLE ALGORITHMS

<Algorithm 1> Initially, allocate a routing function to every area. Then, select the area with the minimum number of packets sent or received per unit time (referred to as “integration-origin area”). If integrating the routing function of the selected area with that of another area (referred to as “integration-destination area”) reduces the total network cost, integrate them. Repeat this until the total network cost can no longer be reduced or there is no integration-origin area left.

<Algorithm 2> Initially, allocate a routing function only to the area where the number of packets sent or received per unit time is the largest. Then, select the area where the number of packets sent or received per unit time is the second largest. If making the routing function of the first area handle the routing of the second area reduces the total network cost, do so. Repeat this until the total network cost can no longer be reduced or there is no area which can be handled by the routing function of another area.

In this paper, we adopt Algorithm 1 and clarify the detailed allocation algorithm.

4.2. DETAILED PROCESSING FLOW OF VIRTUAL ROUTING FUNCTION ALLOCATION

4.2.1. OVERVIEW OF PROCESSING FLOW

The processing flow for determining whether to integrate two areas is described below by referring to Figure 2.

Step 1. Selection of integration-origin area

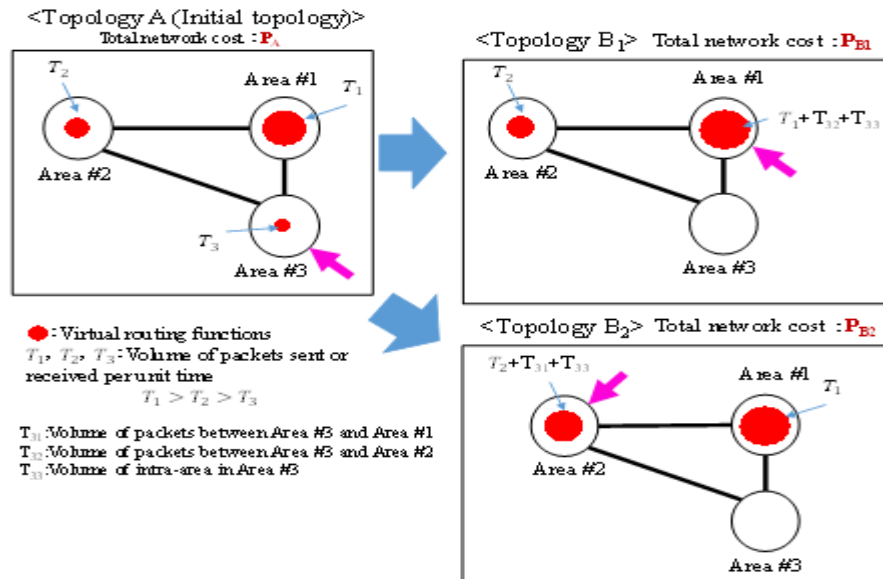
Select an area with the least volume of packets sent or received per unit time, so as to minimize the amount of detour traffic (Area #3 in Figure 2).

Step 2. Selection of integration-destination area

Select all area as integration-destination candidates except integration origin area.

Step 3. Integration

If the total network cost after the integration is smaller than before the integration, the routing function of the integration-origin area is integrated into that of the integration-destination area in which the integration results in the minimum total network cost. For example, if $P_{B1} < P_{B2} < P_A$ in Figure 2, the routing function is integrated into the one in Area #1. Note that P_A , P_{B1} and P_{B2} are total network cost of topology A, B_1 and B_2 respectively. If the total network cost after the integration is not smaller than that before the integration, no integration is made.



(1) Overview of the integration of virtual routing function

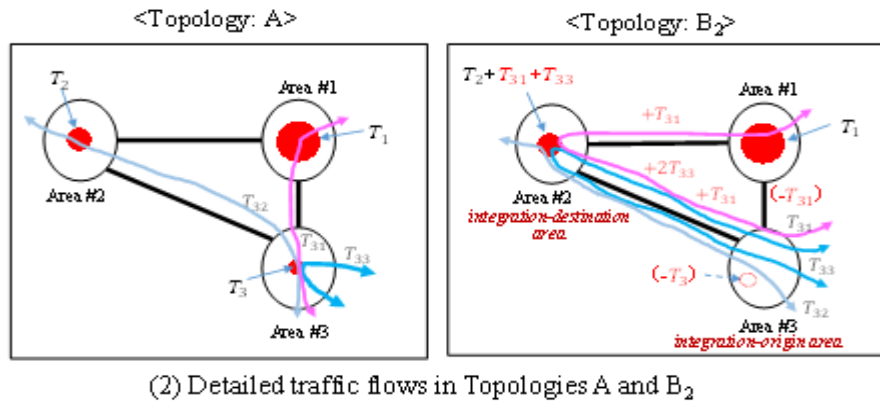


Figure 2. Detailed integration algorithm of the virtual routing function of one area into that of another area

4.2.2 Allocation Policy For Selecting Integration-Source Area In Step 1

Unless the volumes of traffic flows in the two directions between two areas are greatly unbalanced, integration-origin area candidates are selected in descending order of the distance from the center of the network, in **Step 2**. For example, in the ladder-shaped model shown in Figure 3 (which simulates the shape of Japan and consists of 10 areas), it is desirable to reduce the total circuit length by integrating areas into area group G₃ (which contains Areas #5 and #6), which is close to the center of the network so that the total network cost can be reduced. Areas #1, #2, #9, and #10 included in group G₁, which is the farthest from the center of the network, are first selected one by one as integration-origin area candidates. Whether to integrate them is determined using the algorithm described in Section 4.2.1. When all the areas in G₁ have been examined, Areas #3, #4, #7, and #8 in G₂ are selected as integration-origin area candidates in Figure 3.

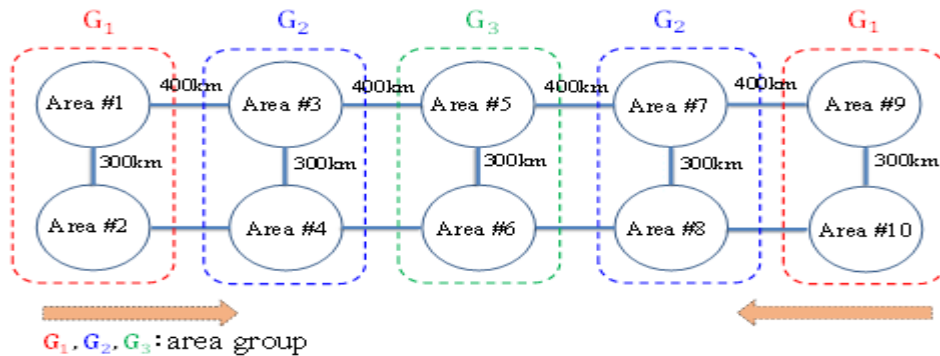


Figure 3. Ladder-shaped network topology which simulates the shape of Japan

5. EVALUATION OF THE PROPOSED ROUTING FUNCTION ALLOCATION

5.1. EVALUATION CONDITIONS

1) Network structure:

The ladder-shaped network illustrated in Figure 3 is used for evaluation. There are 10 areas and the distance between two adjacent areas in the horizontal direction is 400 km while that in the vertical direction is 300 km.

2) Traffic Model

It is supposed that there are 110 simultaneously communicating terminals in each area. If it is assumed that 20% of all terminals are communicating simultaneously, the total number of terminals in all areas will be about 5000 and this number is representative of that of small-to-medium enterprises. The terminals are broken down into 11 sets of 10 terminals, and each set is communicating with a different area. Since it is desirable to consider applications that require the largest bandwidths and the largest number of packets, video traffic is assumed. The bandwidth (one way) required for communication by each terminal is set to 1.8 Mbps, and the packet length is set to 7200 bits. It requires routing capacity of 250 pps (packets/s) per each communication (one way). In addition, to evaluate unbalanced traffic distribution, it is assumed that traffic concentrated on the central area (Area #5 in Figure 3). To evaluate the effect of traffic concentration on Area #5, it is assumed that the volume of traffic between Area 5 and any of the other areas is M times (referred to as “peak coefficient”) larger than that between non-Area #5 areas. If M is one, the traffic distribution is uniform.

3) Route selection between areas

The fixed route is selected between areas, as illustrated in Figure 4. Moreover, it is not considered to select the route dynamically according to the traffic conditions in the network.

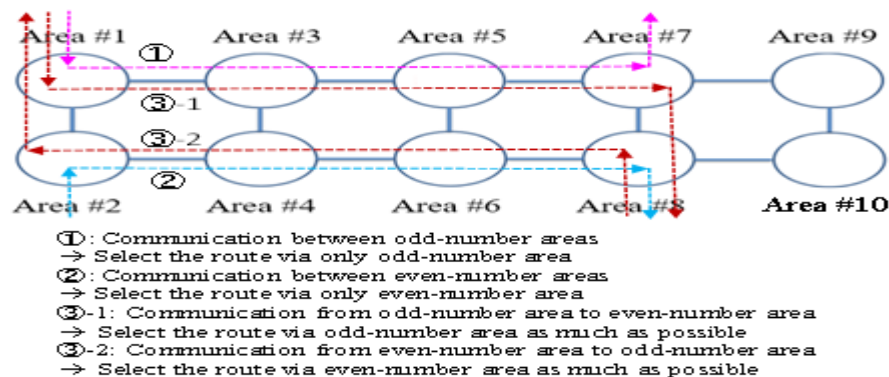


Figure 4. Route selection policies in the evaluation

4) Z_c is defined as the ratio of the cost of a routing function to that of a circuit is critical, as in (1):

$$Z_c = \alpha / \beta \quad (1)$$

where α is the cost of routing functions per pps, and β is the cost of a circuit per Mbps per 10 km. The fixed cost of routing function or circuit bandwidth when there is no traffic is disregarded in this paper.

5) Evaluation tool

We built an evaluation tool written in the C language.

5.2 . RESULTS AND EVALUATION

Evaluation results of the final routing function allocation following the proposed allocation algorithm in Section 4.2 are shown in Figures 5 and 6. The size of a circle, \circ , in an area indicates the capacity of the routing function. Figure 5 shows how the final function allocation changes as Z_c is varied and M is fixed at 2. Figure 6 shows how the final function allocation changes as M is varied and Z_c is fixed at 3. The following points are clear from these Figures:

1) The greater Z_c (cost ratio) is, the greater the cost of the routing function relative to that of the circuit and the more advantageous to integrate the two areas concerned. In the example of Figure 5, routing functions are allocated in all areas when $Z_c \leq 0.9$, in half of all areas when $Z_c = 2.9$, and only in one area when $Z_c > 3.5$.

2) The greater M (peak coefficient) is, the greater the effect of integrating the routing function of one area into that of another area is. This is explained in detail with Figure 7. Two areas, #A and #B are considered in Figure 7. If area #A is integrated into area #B, the circuit capacity that is newly required for intra-area communication in the first area consumes additional cost, x_1 , while cost x_{21} and x_{22} previously consumed by the routing function capacity that had been required to relay traffic to another area becomes zero. The greater M is, the greater x_{22} becomes (even when x_1 remains the same), resulting in an increase in the number of cases where integration is advantageous (the right part in Figure 7).

To ascertain the above-mentioned trend, we examined how the ratio of the number of areas that have routing functions to the total number of areas, δ , changes when Z_c and M are varied. The result is shown in Figure 8. The horizontal axis represents Z_c while the vertical axis represents δ . The minimum value of δ is 0.1 while its maximum value is 1.0. As in Figures 5 and 6, the greater Z_c is and the greater M is, the more advantageous the integration is.

3) Figure 5 shows that, when Z_c is 0.8, the final function allocation is to distribute a routing function to every area. However, this allocation would not be viable in the conventional non-NFV network because a routing function needs to have a capacity that is higher than a certain level, Q . For example, if Q is twice the capacity of the routing function allocated in the case of Figure 5 ($Z_c = 0.9$), the total network cost of the conventional network

(it is the same topology in the case of $Z_c > 3.6$ in Figure 5) increases by 20%. This takes only the cost of the routing function capacity and the circuit bandwidth into consideration, and does not include the fixed cost of servers and circuits, which must be taken into consideration in the conventional network. Therefore, NFV-based routing function allocation could make it possible to distribute small-capacity routing functions in a network, resulting in a large reduction in cost in comparison to the conventional network.

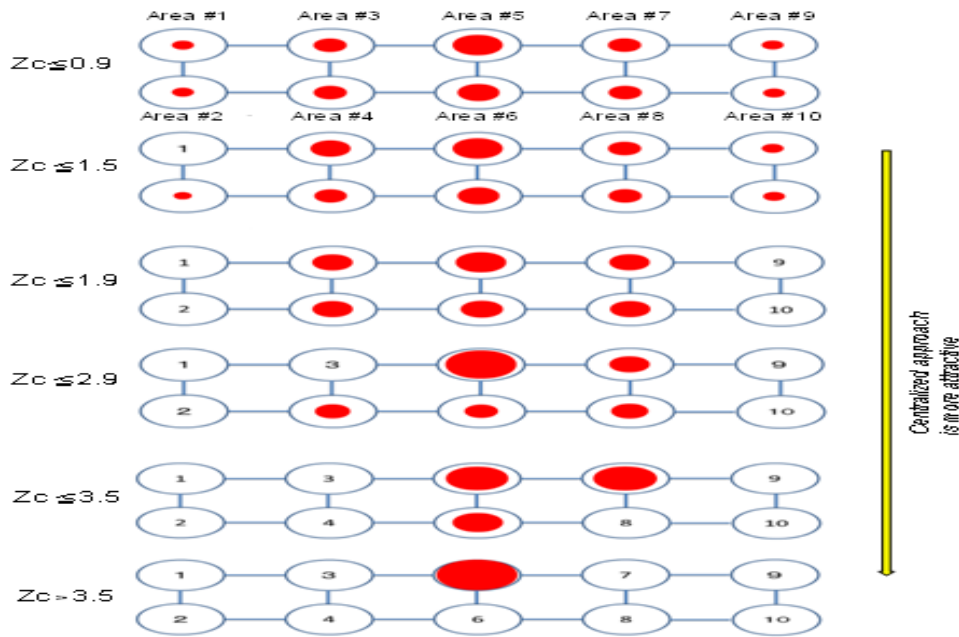


Figure 5. Simulation result 1
 <Final topology of virtual routing function allocation ($M=2$)>

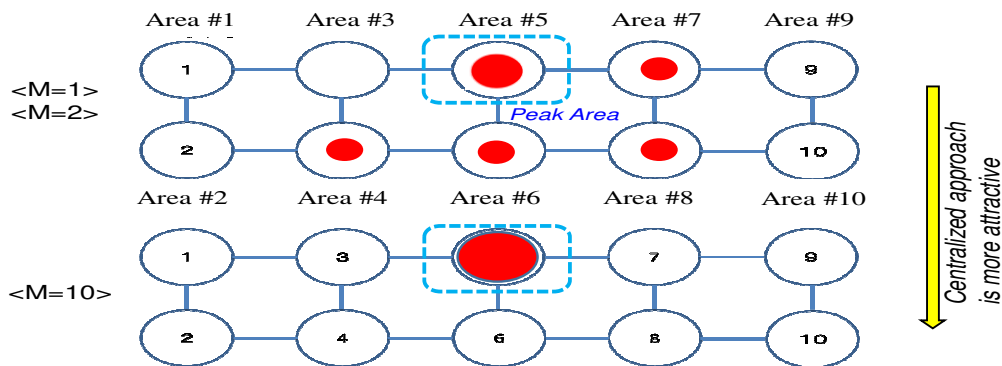


Figure 6. Simulation result 2
 <Final topology of virtual routing function allocation ($Z_c=3$)>

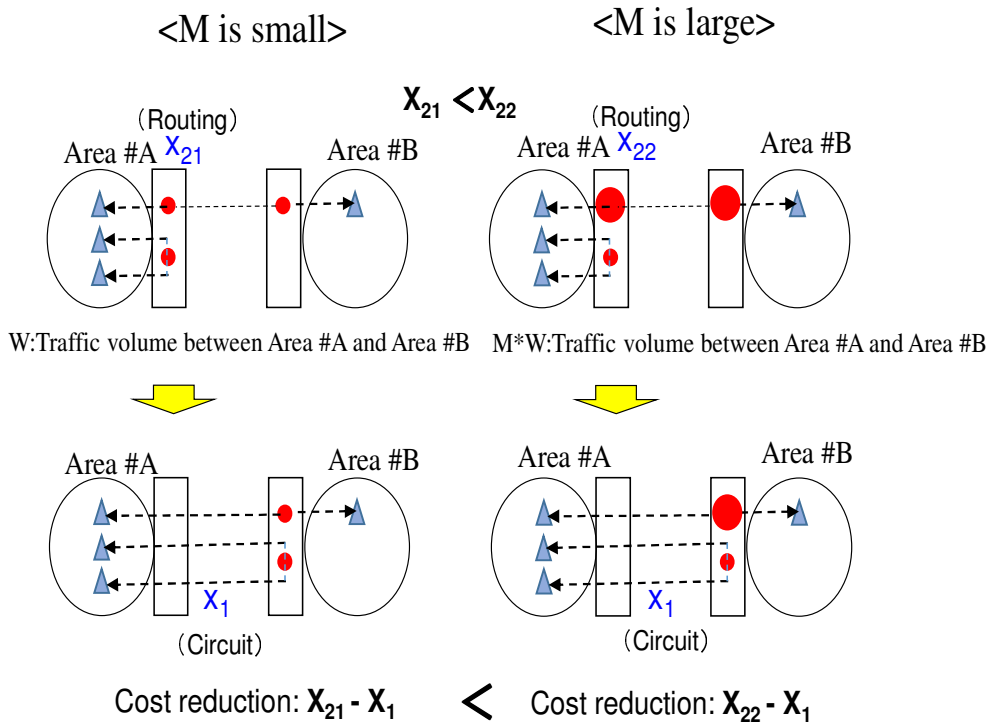


Figure 7. Comparison of area integration effect

5. CONCLUSION

This paper has proposed the algorithm of virtual routing function allocation in the NFV network for minimizing the total network cost. The proposed allocation algorithm has been evaluated with a ladder-shaped network model, which simulates the shape of Japan. It has been found that, the greater the cost of a routing function relative to that of a circuit is and the greater the peak coefficient of traffic between areas is, the smaller the total network cost can be made by integrating the routing functions of some areas into those of a small number of areas. Furthermore, this paper has indicated that there are cases where the use of NFV-based routing functions makes it possible to reduce the total network cost much, compared with a conventional network, in which it is not economically viable to distribute small-capacity routing functions.

It is required to study the routing function allocation, which considers 1) fixed cost of routing functions or circuits where there is no traffic, 2) allowable maximum network delay and 3) dynamic route selection between area. The allocation of other virtual network functions (ex. DPI, FW) in addition to virtual routing functions should also be studied.

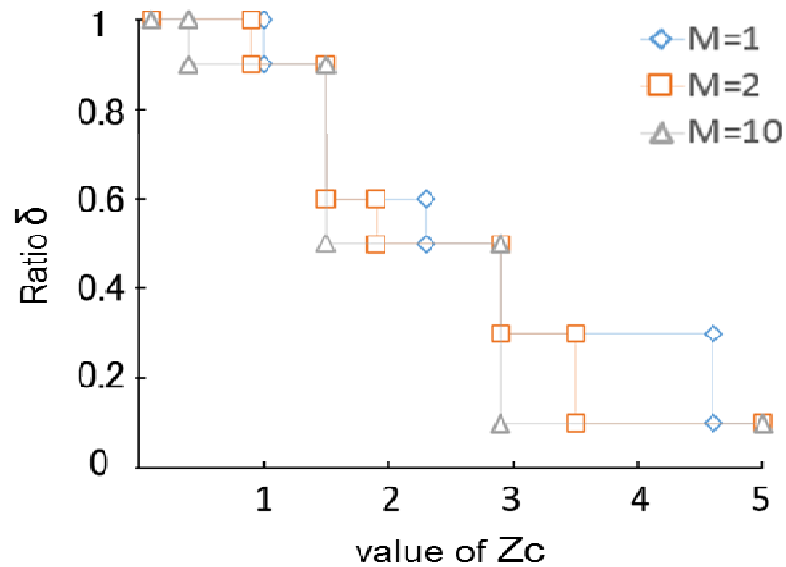


Figure 8. Ratio δ vs. value of Z_c and M

ACKNOWLEDGMENT

I would like to thank Mr. K.Hidafor their help with the simulation.

REFERENCES

- [1] M. Chiosi et al., "Network Functions Virtualisation - An Introduction, Benefits, Enablers, Challenges and Call for Action," ETSI NFV, Oct. 2012.
https://portal.etsi.org/nfv/nfv_white_paper.pdf.
- [2] "Network Functions Virtualisation (NFV); Architectural Framework," ETSI GS NFV 002 v1.2.1, Dec. 2014.
http://www.etsi.org/deliver/etsi_gs/nfv/001_099/002/01.01.01_60/gs_nfv002v010101p.pdf.
- [3] R. Mijumbi, J. Serrat, J. Gorricho, N. Bouten, F. D. Trurck and R. Boutaba, "Network Function Virtualization: State-of-the-art and Research Challenges," IEEE Communications Surveys & Tutorials, Vol. 18, Issue 1, pp.236-262, 2016.
- [4] A. Fischer, J. F. Botero, M. T. Beck, H. Meer, and X. Hesselbach, "Virtual Network Embedding: A Survey", IEEE Communications Surveys & Tutorials, Vol. 15, No.4, 2013.
- [5] X. Wei, S. Hu, H. Li, F. Yang and Y. Jin, "A Survey on Virtual Network Embedding in Cloud Computing Centers", The Open Automation and Control Systems Journal, Vol. 6, pp.414-425, 2014.
- [6] S. Mehraghdam, M. Keller, and H. Karl, "Specifying and placing chains of virtual network functions," in Cloud Networking (CloudNet), 2014 IEEE 3rd International Conference on, Oct 2014, pp. 7-13.

- [7] H. Moens and F. D. Turck, "VNF-P: A model for efficient placement of virtualized network functions," in *Network and Service Management (CNSM)*, 2014 10th International Conference on, Nov 2014, pp. 418–423.
- [8] A. Basta, W. Kellerer, M. Hoffmann, H. J. Morper, and K. Hoffmann, "Applying NFV and SDN to LTE Mobile Core Gateways; The Functions Placement Problem," 4th Workshop on All Things Cellular: Operations, Applications and Challenges 2014, pp.33-38, Aug. 2014.
- [9] M. Bouet, J. Leguay and V. Conan, "Cost-based placement of vDPI functions in NFV infrastructures," *International Journal of Network Management*, pp.490-506, 2015.
- [10] S. Tsumura and S. Kuribayashi, "Simultaneous allocation of multiple resources for computer communications networks", *APCC2006*, 2F-4 (2006.8)
- [11] S. Kuribayashi, "Optimal Joint Multiple Resource Allocation Method for Cloud Computing Environments", *International Journal of Research and Reviews in Computer Science (IJRRCS)*, Vol.2, No.1, pp.1-8, Feb. 2011.
- [12] S. Kuribayashi, "Resource Allocation Method for Cloud Computing Environments with Different Service Quality to Users at Multiple Access", *International Journal of Computer Networks & Communications (IJCNC)* Vol.7, No.6, pp.33-51, Nov.2015.
- [13] R. Mijumbi, J. Serrat, J. Gorricho, N. Boutenz, F. Turckz and S. Davy, "Design and Evaluation of Algorithms for Mapping and Scheduling of Virtual Network Functions," 1st IEEE Conference on Network Softwarization (NetSoft), pp.1-9, 2015.
- [14] S. Clayman, E. Mainiy, A. Galis, A. Manzaliniz and N. Mazzocca, "The Dynamic Placement of Virtual Network Functions," *IEEE Network Operations and Management Symposium (NOMS)*, pp.1-9, 2014.
- [15] K. Hida and S. Kuribayashi, "Virtual Routing Function Allocation Method for Minimizing Total Network Power Consumption," 18th International Conference on Information and Communication Systems (ICICS Venice 2016), Aug. 2016.

AUTHOR

Shin-ichi Kuribayashi received the B.E., M.E., and D.E. degrees from Tohoku University, Japan, in 1978, 1980, and 1988 respectively. He joined NTT Electrical Communications Labs in 1980. He has been engaged in the design and development of DDX and ISDN packet switching, ATM, PHS, and IMT 2000 and IP-VPN systems. He researched distributed communication systems at Stanford University from December 1988 through December 1989. He participated in international standardization on ATM signaling and IMT2000 signaling protocols at ITU-T SG11 from 1990 through 2000. Since April 2004, he has been a Professor in the Department of Computer and Information Science, Faculty of Science and Technology, Seikei University. His research interests include optimal resource management, QoS control, traffic control for cloud computing environments and green network. He is a member of IEEE, IEICE and IPSJ.

