3D On-chip Data Center Networks Using Circuit Switches and Packet Switches

Takahide Ikeda Yuichi Ohsita, and Masayuki Murata
Graduate School of Information Science and Technology, Osaka University
Osaka, Japan
{t-ikeda, y-ohsita, murata}@ist.osaka-u.ac.jp

***Abstract*—The energy consumption of the data center becomes a great problem. One approach to reduce the energy consumption of the data center is to use *on-chip data centers*, which are integrated circuit chips that perform the tasks in a data center. On-chip data centers are constructed of cores and the network between cores. Because the tasks in the data center are performed by the cooperation between servers, the network between cores in the on-chip data center may have a large impact on the performance of the chip. In this paper, we investigate the network structures for the on-chip data centers. We focus on the 3Dnetwork using both circuit and packet switches, and compare the energy consumption and the delay of the candidate network structures. The results show that (1) the servers should connect to the packet switches in the same layer, (2) the packet switches should connect to the circuit switches in all layers, and (3) the layer including both of circuit switches and packet switches should be avoided to reduce the energy consumption and the delay.**

***Keywords*—network on chip; data center; energy consumption; delay; 3D on-chip network**

I. INTRODUCTION

In recent years, online services such as cloud Computing have become popular, and the amount of data, required to be processed by such online services is increasing. Such a large amount of data is handed by data centers, and many data centers have been built. As the services provided by data centers become popular, the energy consumption of the data center becomes an important problem. The energy consumed by data centers occupies 1.5% of the total energy consumption consumed in the world [1]. Thus, an energy efficient data center is required.

A data center is constructed of many servers. In a data center, each server performs its assigned task, cooperating with other servers [2]. The data center can process a large amount of data because a server cooperates.

One approach to reduce the energy consumption caused by the data center is to make an integrated circuit chip that can perform the above tasks in a data center. This kind of chip is called an *on-chip data center* [3]. An on-chip data center is made of a large number of CPU cores and the network between the cores on a single chip. An on-chip data center works with a significantly small energy because of its small wiring length of the network within a chip [4].

Most of existing work on on-chip data centers focus on the usage of many cores on the chip. However, because tasks in a data center require communication between servers, the network structures between cores may have a large impact on the performance and/or the energy consumption of the on-chip data center.

The network within a chip is often called a *Network on chip (NoC)*, and constructed of switches. Two types of switches are used in a NoC: packet switches and circuit switches.

A packet switch relays packets or flits, which are a small pieces of a packet, based on their destination addresses. On the other hand, a circuit switch connects its input port with one of its output ports based on the configuration. A circuit switch consumes a small energy compared with a packet switch because it does not require any processing to relay traffic, though multiple flows from different input ports cannot share the same output port.

Several NoC architectures that use both packet and circuit switches have been proposed [5-7]. In these architectures, the circuit path between packet switches is established by configuring the circuit switches along the route of the circuit path. The set of the packet switches and the established circuit paths constructs the network topology. In these architectures, the network topology can be changed by the configuration of the circuit switches. Stensgaard et al. [7] proposed a method to configure the circuit switches suitable to the application before starting the application.

The network architectures using both of packet and circuit switches are also effective in an on-chip data center. In a data center, though the traffic pattern changes significantly and frequently, each server communicate with only a small number of servers at once [8]. Considering such traffic, the network topology where the communicating server pairs are connected closely is preferable. This network topology can be set by setting the circuit switches in the network using both of the packet and circuit switches. Even if the traffic pattern changes, we change the network topology so as to suit the current traffic pattern by reconfiguring the circuit switches.

In recent years, another new NoC architecture called *3D NoC* has been proposed [9]. The 3D NoC is constructed by stacking multiple 2D chip layers vertically. The vertically stacked layers decrease the number of hops between switches. Moreover, the vertical links of the 3D NoC are significantly shorter than the horizontal links. As a result, the 3D NoC reduces both the energy consumption and the delay.

In addition, the 3D NoC improves the effectiveness of using packet and circuit switches. Because the 3D NoC increases the number of candidate routes of the circuit paths, we establish more circuit paths, which reduce both the energy consumption and the delay. However, the 3D NoC using both packet and circuit switches has not been discussed sufficiently.

In this paper, we investigate the network structures suitable for the on-chip data center. In an on-chip data center, a server is constructed by multiple directly connected cores. Then,

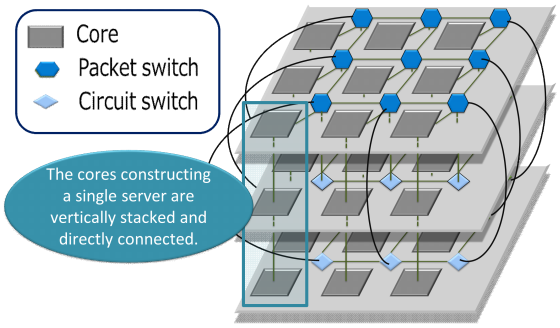


Figure. 1. 3D on-chip data center network

the network connects the servers. In this paper, the network between servers is constructed as a 3D network using circuit and packet switches. We investigate the network structures, focusing on the following three points; (1) connection between layers in the 3D network, (2) connection between servers and switches, and (3) placement of switches within each layer. The results show that (1) all servers should be connected to the packet switches in the same layer, (2) all packet switches should be connected to all layers, and (3) each layer should have only the same type of switches.

The rest of this paper is organized as follows. Section II explains the overview of the on-chip data center used in this paper. In Section III, we investigate the network structures suitable to the on-chip data center. Section IV presents the conclusion.

II. On Chip Data Center Network

A data center is constructed of servers and a network between servers. The tasks in a data center, such as handling a large amount of data, are performed by servers cooperating with each other. Such a task in a data center is split into subtasks, and each subtask is assigned to and performed by one of the servers. Each server obtains the data or the results of the other subtasks from the other servers, if the data or the results are required to complete its subtask.

In this paper, we investigate the on-chip data center, which performs tasks in a data center. The on-chip data center used in this paper is constructed of multiple cores and a network between cores. The tasks are handled by multiple cores in a data center. Each of the cores in the on-chip data center provides a resource or cache memory. The related multiple cores are connected to each other, and act as a single server in a data center. We call these connected multiple cores a *server* in the on-chip data center. The network structure used in this paper is shown in Fig 1. In this structure, the cores constructing a single server are vertically stacked and directly connected. Then, the servers are placed in a lattice.

The network between servers is constructed of switches placed in a 3D lattice, because the lattice network can be easily constructed on a chip. Each server is connected to the network by connecting one of its core to one of the switches.

In the on-chip data center, we use two kinds switches, i.e., packet switches and circuit switches. The packet switches and the circuit switches have their advantages and disadvantages. The circuit switches consume less energy than the packet switches, because the circuit switches do not require complicated processing such as decision of the next hop. However, the circuit switch cannot relay flows from different input ports to the same output port. On the other hand, the flows from the different input ports share the same output port in the packet switch, though the packet switch consumes more energy.

In this paper, we use both types of switches as follows. All servers are connected to packet switches, so that each server can communicate with multiple servers at once. The switches not connected to servers are circuit switches because the circuit switches consumes less energy. In this network, the traffic is sent after constructing the network topology by setting the circuit paths between packet switches. The circuit paths are established by configuring the circuit switches along the paths. Then, the traffic is sent over the network topology of the packet switches constructed by the circuit paths.

This network structure has the following parameters; (1) the connection between layers, (2) the layers where switches connected to servers are deployed, and (3) the types of switches deployed in each layer, which are discussed in Section III.

III. Comparison of on-Chip Data Center Networks

A. Network structures

In this section, we investigate the following parameters of the 3D network structures for on-chip data center.

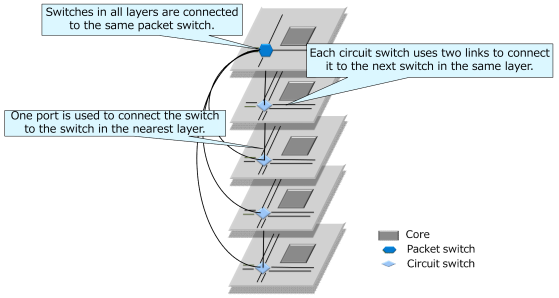
1) Inter-Layer Connection: There are two types of the inter-layer connection. The first type is shown in Fig 2(a). In this case, switches in all layers are connected to the same packet switch. We call this type of connection the *packet switch centric connection*.

Another type of the inter-layer connection is shown in Fig 2(b). In this case, all vertical links are constructed only between nearest layers. We call this type of connection the *nearest layer connection*.

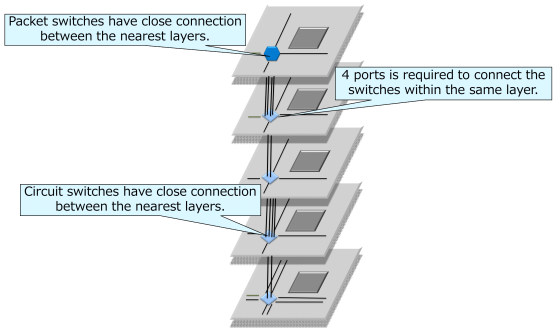
In our comparison described in Subsection III-B1, we deploy all packet switches at the first layer. All packet switches have 8 ports and all circuit switches have 10 ports in both types of connections. We set the number of layers to 5. In the circuit switch centric connection, each circuit switch uses two links to connect it to the next switch in the same layer. That is, each circuit switch uses 8 ports to the connection in the same layer. Two ports per circuit switch are used to the vertical connection. One of them is used to connect the switch to the packet switch at the first layer. The other port is used to connect the switch to the switch in the nearest layer.

In the nearest layer connection, we add close connection between the nearest layers. All vertical links from the packet switches at the first layer are connected to the switches at the second layer. Thus, 4 ports of the switches at the second layer are connected to the switches at the first layer. Among the residual ports, 4 ports are required to connect the switches within the same layer. Finally, the other ports are used to connect the switches to the third layer. The switches in the other layers are connected in a similar way.

The comparison of these connections clarifies whether the one hop connection from the packet switches to any layers is preferable or the close connection between the nearest layer is preferable.



(a) Packet switch centric connection



(b) Nearest layer connection

Figure. 2. Inter layer connection

2) Layer of Swithes Connected to Servers: In the on-chip data center investigated in this paper, each server is connected to one of the switches nearest to the server. As shown in Fig 3, there are two types of connections between servers and switches. In the first type of connection, all servers are connected to the switches in the same layer. We call this type of connection the *same layer connection*. In the other type of connection, the servers neighboring with each other are connected to the switches in the different layers.

In the same layer connection, the number of hops between servers is small because all servers are connected in the same layer. However, the connections of packet switches at the first layer are static. On the other hand, the connections between

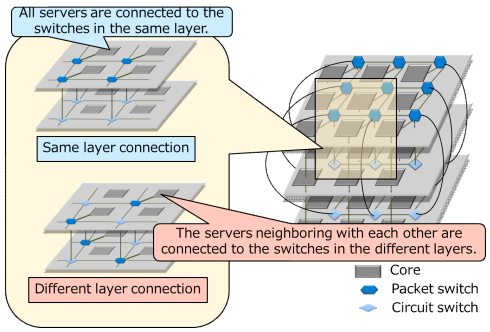
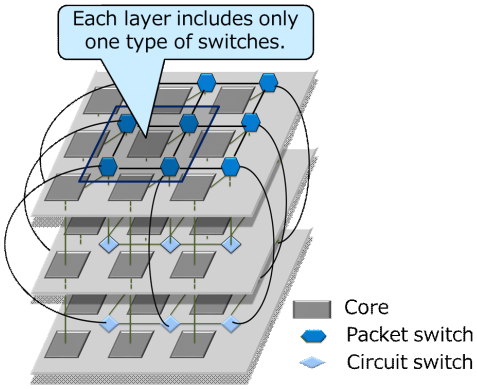
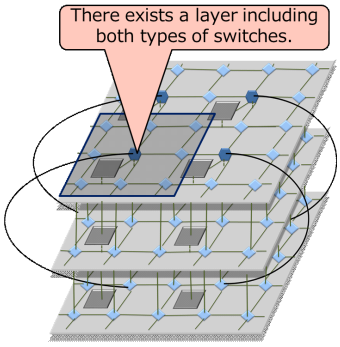


Figure. 3. Connection from servers



(a) Single type placement



(b) Multiple type placement

Figure. 4. Placement of switches within each layer

packet switches can be changed in any layers in the different layer connection.

3) Placement of Switches within a Layer: There are two kinds of placement of the switches in the same layer. The first one is shown in Fig 4.(a). In this type of the placement, each layer includes only one type of switches. We call this type of the placement the *single type placement*. In the other type of placement, there exists a layer including both types of switches. We call this type of placement the *multiple type placement*.

The *multiple type placement* has more candidates of routes of circuit paths between the packet switches than the *single type placement*. Thus, the energy efficient routes may be found, even when the number of flows to be accommodated is large. However, the number of switches passed by each flow between servers increases. On the other hand, the *single type placement* has less routes between the server pairs. However, the numbers of switches passed by each flow between servers are small.

B. Models Used in Our Comparison

1) Energy consumption model: The energy consumed by the network on chip depends on (1) network structure, (2) the traffic amount on the network, and (3) the bit flips of the traffic.

Wolkotte et al. [10] model the energy consumed by a circuit switch, a packet switch and a link in the NoC, in the case of 50% bit flips. In this model, the circuit switch consumes 0.37 μW, the packet switch consumes 0.98 μW, and the link consumes $(0.39 + 0.12L)$ μW where L is a length of link

(mm) to relay 1 bit of traffic. In this paper, we use this model to evaluate the energy consumption. In this paper, we focus only on the energy consumed by the network, and exclude the energy consumed by the cores.

Though the actual energy consumed by the switches and links may be different from this model, the packet switch consumes more energy than the circuit switch in any cases because the packet switch requires more process such as checking the destination of each packet. Therefore, the results in this paper are independent of the switch architectures.

2) Delay Model: In this paper, we also compare the delay between cores. We define the delay as the time required to receive all traffic by the destination cores after generating the traffic demands.

In the network on chip, packets are generally divided into flits, and each switch relays the flits. In this paper, we assume that each flit can be relayed by a packet switch to the next packet switch in 1 clock cycle. Though, the clock cycle required to relay a flit depends on the switch architectures and may be different from this model. The suitable network structures discussed in this paper are independent of switch architectures because the order of delays is the same as the results in this paper even if multiple clock cycles are required to relay a flit.

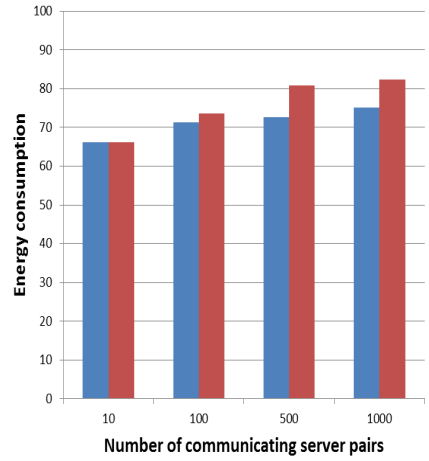
In the on-chip data center, we also use the circuit switches. The circuit switch is configured to connect the input and output ports in advance. The packet switches can be connected by configuring the circuit switches. The packet switch pairs, connected by the circuit paths, relay the flits by the same way as the packet switches which are directly connected to each other. The relay of the flits by the circuit switch takes no clock cycles. Thus, the delay between cores depends only on the number of packet switches passed by the flow.

3) Traffic Model: According to Benson et al. [8], each server communicates with only a small number of servers at once. Therefore, we generate traffic between randomly selected server pairs. In our evaluation, we vary the number of communicating server pairs from 10 to 1000. The traffic rates between selected server pairs are set to 10000.

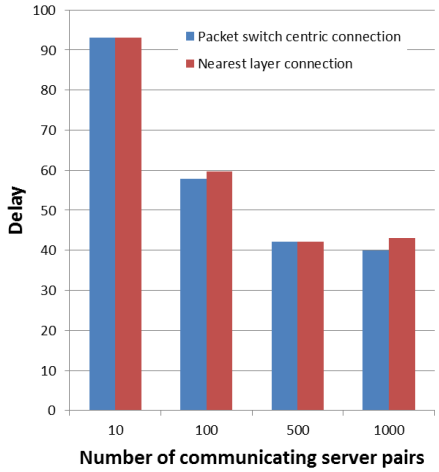
4) Path Computation Model: We calculate the routes of traffic so as to make the energy consumed by the traffic small.

In this paper, we calculate the route of all traffic demands. The route of each traffic demand is calculated by the Dijkstra algorithm setting the weights of the links to the energy consumed to relay the traffic. If the calculated route uses the circuit switch, we connect both ends of the input and the output ports, and remove the ports of the circuit switch before the calculation of the routes of the next traffic demands, so as to avoid the output ports of the circuit switch used by the other traffic from the different input ports.

In this path computation, we assume that the traffic demands are known before calculating the routes. By using this model, we discuss the suitable network structure when the routes are calculated optimally. However, the actual traffic demands may be unknown when calculating routes, and we require a method to calculate the routes without traffic demand information, which is one of our future work.



(a) energy consumption



(b) Delay

Figure. 5. Comparison of inter-layer connections

C. Results

In our evaluation, we use the network structure with 5 layers and 255 servers. We generate 4 patterns of traffic, and compare the average of the energy consumption and the delay.

1) Comparison of Inter-layer connections: In this subsection, we compare the network structures of different inter-later connections. The comparison of energy consumption is shown in Fig 5(a). The vertical axis of the figure indicates the energy consumption normalized so that the energy consumption in the 2D lattice using only the same number of packet switches as the 3D network structures used in this comparison becomes 100. Fig 5(a) shows that both of the inter-layer connections reduce the energy consumption compared with the 2D lattice. This is because the 3D lattice structures used in this comparison establish the circuit paths to reduce the energy consumption. However, the energy consumption in both types of the connections becomes close to that of the 2D lattice when the number of communicating server pair becomes large. This is because we cannot establish energy efficient circuit paths for all communicating server pairs. As a result, a large amount of traffic passes multiple packet switches similar to the 2D lattice.

Fig 5(a) also shows that the energy consumption of the packet switch centric connection is smaller than the nearest layer connection. This is because a flow is required to pass multiple layers to use the circuit switch whose layer is far from the packet switch in the nearest layer connection. Because each switch relaying the traffic consumes energy, the large number of switches passed by each flow cause a large energy consumption. On the other hand, the packet switches are directly connected to all layers in the packet switch centric connection. Thus, the number of switches passed by traffic is smaller than the nearest layer connection. As a result, the packet switch centric connection accommodates traffic with a smaller energy consumption than the nearest layer connection.

The comparison of delay is shown in Fig 5(b). The vertical axis of the figure indicates the delay normalized so that the delay in the 2D lattice using only packet switches becomes 100. Fig 5(b) shows that the 3D network structures using both of circuit switches and packet switches reduce delay significantly compared with the 2D lattice. This is because the circuit paths reduce the number of hops of packet switches.

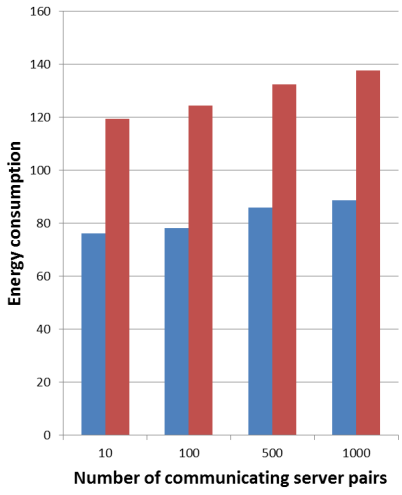
The normalized delay becomes small when the number of communicating server pairs increases. This is because the circuit paths balance the loads. In the case of the 2D lattice, traffic concentrates at some packet switches, and is required to wait to be relayed, when the number of communicating server pairs becomes large. On the other hand, in the 3D lattice using circuit switches and packet switches, the circuit paths directly connect the packet switches which are far from each other, and avoid concentration of traffic on a certain switch.

Fig 5(b) shows that the packet switch centric connection and the nearest layer connection achieve the similar delay. This is because the circuit switches do not have an impact on the delay though the traffic passes more circuit switches in the nearest layer connection than the packet switch centric connection.

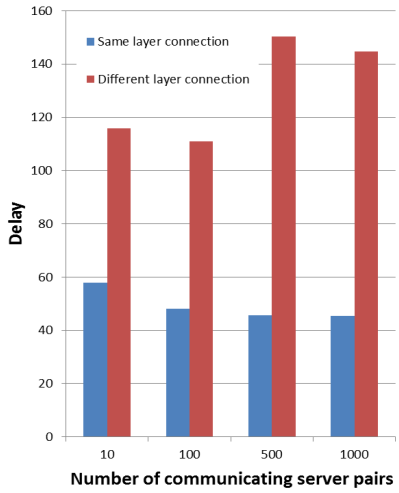
2) Comparison of the Connection between Servers and Switches: In this subsection, we compare the network structures of the different types of the connections between servers and switches.

The comparison of energy consumption is shown in Fig 6(a). The vertical axis of the figure indicates the energy consumption normalized so that the energy consumption in the 2D lattice using only packet switches becomes 100. Fig 6(a) shows that the same layer connection achieves the lower energy consumption than the 2D lattice. On the other hand, the different layer connection consumes more energy than the 2D lattice. This is because the traffic between servers passes more switches in the different layer connection than the 2D lattice.

The comparison of delay is shown in Fig 6(b). The vertical axis of the figure indicates the delay normalized so that the delay in the 2D lattice using only packet switches becomes 100. Fig 6(b) shows that the delay in the different layer connection becomes larger than the 2D lattice. This is because of the large number of hops of packet switches. In the different layer connection, the packet switches exist not only at the first layer, but also the other layers. Such packet switches block the long circuit path, and even cause a large number of packet switches passed by a flow. On the other hand, the same layer



(a) Energy consumption



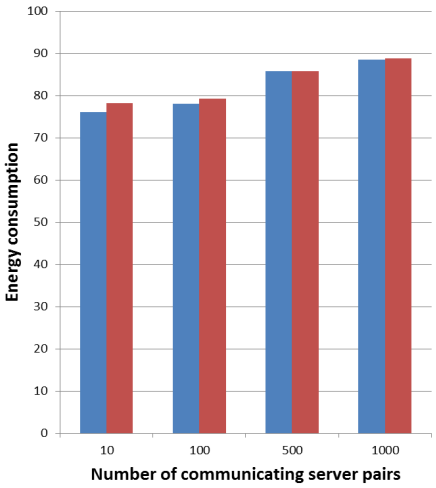
(b) Delay

Figure. 6. Comparison of the connection from servers

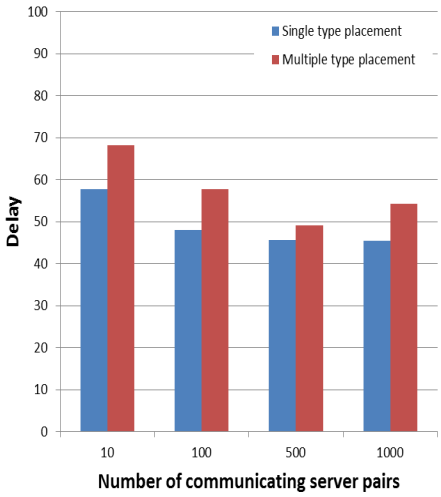
connection reduces the delay significantly compared with the 2D lattice. This is because long circuit paths are established in the same layer connection, and reduce the number of packet switches passed by a flow.

Similar to the results in Fig 5(b), Fig 6(b) also indicates that the normalized delay becomes small in the same layer connection when the number of communicating server pairs becomes large. This is because the circuit paths balance the loads. On the other hand, the normalized delay becomes large in the different layer connection. In the different layer connection, we cannot add the long circuit paths. Moreover, the number of packet switches passed by a flow is large. As a result, traffic concentrates at some packet switches.

3) Comparison of Placement of Switches within a Layer: In this subsection, we compare the types of the switches used in each layer. The multiple type placement increases the number of candidate routes for the circuit paths. However, the number of hops becomes larger than the single type connection. Comparing them, we clarify whether the larger



(a) Energy consumption



(b) Delay

Figure. 7. Comparison of the placement of switches within each layer

number of candidate circuit paths is preferable or the smaller number of hops between servers is preferable.

The comparison of energy consumption is shown in Fig 7(a). The vertical axis of the figure indicates the energy consumption normalized so that the energy consumption in the 2D lattice using only packet switches becomes 100. Fig 7(a) shows that the single type placement and the multiple type placement consume the similar energy. The multiple type placement has more routes of the circuit paths between servers than the single type placement, and can find energy efficient routes. However, the number of hops between servers becomes large, which consumes more energy. In the case of our simulation, the amount of the energy reduced by using circuit paths equals the amount of the energy increased by the increase of switches relaying the traffic.

Similar to Fig 5(a), Fig 7(a) also shows that both types of the placements consume the similar energy to the 2D lattice using only packet switches when the number of communicating servers becomes large. This is because we cannot set the circuit paths for all communicating server pairs.

The comparison of the delays is shown in Fig 7(b). The vertical axis of the figure indicates the delay normalized so that the delay in the 2D lattice using only packet switches becomes 100. In Fig 7(b), the single type placement achieves the smaller delay than the multiple type placement. This is because the routes are set so as to make the energy consumption small. Though the multiple type placement has more candidates routes for the circuit paths, our route calculation selects the routes that make the energy consumption small, even if the routes cause the concentration of traffic.

In addition, Fig. 7(b) also indicates that the normalized delay becomes small when the number of communicating server pairs becomes large similar to Fig. 5(b) and Fig. 6(b).

IV. Conclusion and Future Work

In this paper, we evaluated the 3D on-chip network structures for the on-chip data centers, which uses both of the circuit and packet switches. According to the results, to reduce the energy consumption and delay, (1) the servers should connect to the packet switches in the same layer, (2) the packet switches should connect to the circuit switches in all layers, and (3) the layer including both of circuit switches and packet switches should be avoided.

Our future work includes the method to calculate the routes suitable to the on-chip networks.

Acknowledgments

This work was partially supported by JSPS Grant-in-Aid for Young Scientists (B)23700077.

References

[1] J. G. Koomey and P. D, "Growth in data center electricity use 2005 to 2010," *The New York Times*, Aug. 2011.

[2] D. Abts and B. Felderman, "A guided tour of data-center networking," in *Communications of the ACM*, vol. 10, Jun. 2012, pp. 44–51.

[3] M. Kas, "Toward on-chip datacenters: a perspective on general trends and on-chip particulars," in *The Journal of Supercomputing*, vol. 62, Oct. 2012, pp. 214–226.

[4] R. Iyer, R. Illikkal, L. Zhao, S. Makineni, D. Newell, J. Moses, and P. Apparao, "Datacenter-on-chip architectures: Tera-scale opportunities and challenges in Intel's manufacturing environment," in *Intel Technology Journal*, vol. 11, Aug. 2007, pp. 227–237.

[5] T. Bjerregaard and S. Mahadevan, "A survey of research and practice of network-on-chip," vol. 1-51, Mar. 2006.

[6] M. B. Stensgaard and J. Sparso, "ReNoC: A network-on-chip architecture with reconfigurable topology," in *Proceedings of the Second ACM/IEEE International Symposium on Networks-on-Chip*, Apr. 2008, pp. 55–64.

[7] M. Modarressi, H. Sarbazi-Azad, and M. Arjomand, "A hybrid packet-circuit switched on-chip network based on sdm," in *Proceedings of the Design, Automation & Test in Europe Conference & Exhibition, 2009. DATE '09*, Apr. 2009, pp. 566–569.

[8] T. Benson, A. Anand, A. Akella, and M. Zhang, "MicroTE : Fine grained traffic engineering for data centers," in *Proceedings of the Seventh COnference on emerging Networking EXperiments and Technologies*, Dec. 2011, pp. 1–12.

[9] F. Li, C. Nicopoulos, T. Richardson, and Y. Xie, "Design and management of 3D multiprocessors using network-in-memory," in *Proceedings of ISCA*, Jun. 2006, pp. 130–141.

[10] P. T. Wolkotte, G. J. M. Smit, N. Kavaldjiev, J. E. Becker, and J. Becker, "Energy model of networks-on-chip and a bus," in *Proceedings of IEEE International Symposium on System-on-Chip*, Nov. 2005, pp. 82–85.