

Energy saving through a buffer control approach for hybrid optoelectronic routers

Yuichi Ohsita, Takamichi Nishijima, Yuki Koizumi, and Masayuki Murata

Abstract—Data center networks are required to have high energy efficiency as well as high communication performance. One approach to achieving these requirements is to use a hybrid optoelectronic router, which has optical packet switching functionality and packet buffering in an electronic domain. These hybrid optoelectronic routers can cut the power of electronic buffers while maintaining optical packet switching functionality. This enables energy consumption to be reduced while maintaining the required communication performance. The routers whose buffers can be shut down depend on the traffic routes; more buffers can be shut down by setting the traffic routes so as to avoid packet collision. However, the routes that avoid packet collision may increase the number of powered-on routers. Therefore, we should select the traffic routes considering the energy consumptions of the routers themselves and their buffers. In this paper, we propose a heuristic route selection method to reduce energy consumption while maintaining the required communication performance. For high energy efficiency, our method efficiently re-uses routers and buffers without degrading the required communication performance. Through simulation, we demonstrate that our method reduces the energy consumption significantly compared to methods that do not consider buffer energy consumption or that calculate the shortest traffic routes through the network.

Index Terms—energy saving; data center network; hybrid optoelectronic router; route selection; reuse easiness;

I. INTRODUCTION

A data center network plays an important role in a data center, and is hence required to have high communication performance. In a data center, a large amount of data is handled by many servers cooperating with each other. A lack of bandwidth or large delay prevents communication between servers and increases the time taken to obtain the required data. This degrades the performance of the data center.

Another serious problem in data centers is energy consumption. The energy consumed by data centers increases as the amounts of data they handle rise. The energy consumption of the network occupies a non-negligible percentage of the total energy consumed in the data center [1]. Therefore, data center networks that have high energy efficiency and high communication performance are necessary [2].

Optical networking is a promising solution for future networks with high energy efficiency and communication performance [3]. Optical network devices provide low latency

communication and small energy consumption because they relay optical signals without any conversion to electrical signals. Optical networking also provides a large bandwidth because of technologies such as wavelength division multiplexing (WDM).

There are several studies on data center networks that use optical network devices. One study presented a data center network using optical path switches [4]. In this architecture, optical path switches are placed at the core of the data center network. They hence provide a large bandwidth between server rack pairs transferring huge amounts of data by establishing optical paths. However, this architecture cannot handle frequent traffic changes because it takes time to establish the optical paths.

Another approach to using optical devices in a data center network is to use optical packet switches [5]–[7]. Optical packet switches relay optical packets based on the optical labels attached to the optical packets. Because optical packet switches do not require the establishment of paths, a network constructed of optical packet switches can handle frequent changes in traffic.

In recent years, data center architectures using optical packet switches have been proposed. Chao and Xi have built bufferless optical packet switches for a data center [8]. In this packet switch, the timing to send a packet must be controlled by a scheduler to avoid collisions. Deploying a buffer at each optical packet switch avoids collisions without scheduling. However, this optical buffer is still under development, and the current optical buffer has only a small capacity [9].

To overcome the above problems, a hybrid optoelectronic router has been proposed [10]–[12]. The hybrid optoelectronic router has an optical packet switching functionality with packet buffering in the electronic domain. Unless packet collision occurs on a router, the router can directly relay the optical packets without conversion to electrical signals. Even if a collision occurs, the router can retransfer the packets after storing them in the electronic buffer.

The hybrid optoelectronic routers can cut power to electronic buffers while keeping optical packet switching functionality. This enables them to reduce energy consumption while maintaining the required communication performance.

One approach to reducing the power of electronic buffers is to turn on buffers only when they become required; buffers in the routers where collisions are detected are turned on, and buffers that are not used are turned off. However, in this approach, turning on buffers may take time, and packets can be dropped before the buffer becomes ready.

Another approach is to turn off only the buffers that can

Manuscript received XX XX, 20XX

This work was done when all the authors were with the Graduate School of Information Science and Technology, Osaka University, Suita-shi, 565-0871, Japan (e-mail: y-ohsita, ykoizumi, murata@ist.osaka-u.ac.jp).

This work was funded by the National Institute of Information and Communications Technology (NICT) R&D program, “Basic technologies for high-performance optoelectronic hybrid packet router”

never be used. This approach avoids packet loss unless the buffer size becomes too small by turning on the buffers of routers where collisions are possible in advance. The buffers of routers for which collisions are possible depend on the traffic routes; collision occurs at the routers where incoming traffic from different input ports is directed to the same output port. Therefore, this approach should be combined with a method to set the routes; the routes are configured so as to minimize the energy consumption of the buffers required to be turned on. The unnecessary routers and buffers are then turned off.

In this paper, we propose a method to configure the routes so as to minimize this energy consumption. In this method, we deploy a controller that configures all routes within a data center. The routing controller periodically collects the traffic demands and calculates routes that minimize the energy consumption under the constraint that all traffic demands can be accommodated. To calculate the routes, we propose a heuristic method that calculates the routes of traffic flows between server racks one-by-one. When calculating a route for each flow, we construct a graph whose vertices correspond to the links of the data center network. In this graph, the vertices corresponding to the links connected to the same router are connected by an edge. We assign a weight to each edge based on the additional energy required to use the corresponding router, considering the energy consumption of optical packet switching and buffers. Using this graph, we may calculate the routes with the minimum energy consumption by simply calculating the shortest routes on the graph.

We evaluate our method by comparing it with a method that minimizes the number of powered-on routers without considering buffer energy consumption as well as a method that calculates the shortest traffic routes through the network. The results demonstrate that our method saves energy consumption significantly by considering buffer energy consumption as well as overall network energy consumption.

The construction of this paper is as follows. Section II summarizes previous work related to energy savings on networks. Section III gives an overview of a hybrid optoelectronic routers and a data center network that uses these routers. Section IV presents an overview of our energy saving strategy. Section V proposes the method to calculate routes that minimize energy consumption, and Section VI presents our evaluation. Finally, Section VII summarizes this paper and discusses directions for future study.

II. RELATED WORK

A number of studies have reported that turning off unnecessary devices achieves high energy efficiency in networks [13]–[21].

Wang *et al.* presented an analysis of both VM assignment and network routing with respect to energy conservation [14]. Kilazovich *et al.* proposed a scheduling approach that combines energy efficiency and network awareness [15]. Mingui *et al.* proposed a traffic engineering method to reduce energy consumption on a core network with IP routers [16]. Heller *et al.* proposed a network-wide control method to turn off switches and links on the basis of current traffic [17].

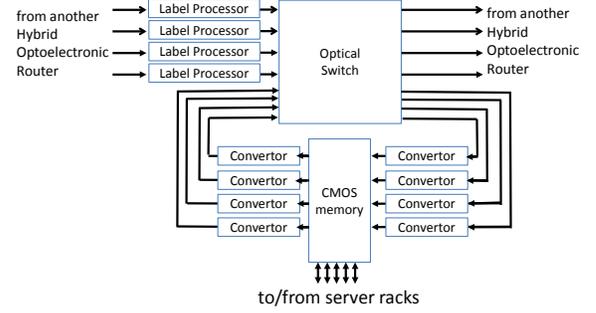


Fig. 1. Hybrid optoelectronic router

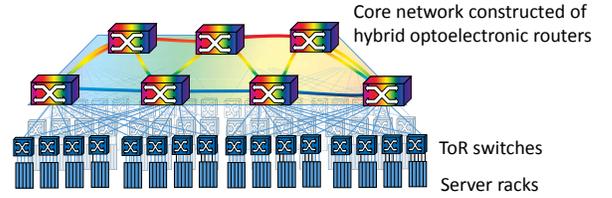


Fig. 2. Network with hybrid optoelectronic routers in a data center

In the above papers, the energy consumption is reduced by minimizing the numbers of powered-on routers or powered-on ports. The number of powered-on buffers is not considered in these papers. However, in a network constructed of hybrid optoelectronic routers, considering buffer energy consumption may reduce energy consumption more efficiently because a hybrid optoelectronic router can transfer packets, even if its buffer is powered off. Therefore, in this paper, we discuss a strategy to reduce the energy consumption of a network constructed of hybrid optoelectronic routers, considering the energy consumption saved by shutting down the buffers.

For a translucent optical network, the placement of the optical-to-electrical or electrical-to-optical converters and regenerators considering the energy consumption has also been discussed [22]–[24]. The converters and regenerators are necessary to keep the signal quality high for long-distance communication, but consume a large amount of energy. Therefore, these methods minimize the number of converters and regenerators under the constraint that the signal quality is kept sufficiently high.

In a data center network, the signal quality does not degrade significantly because all devices are in a single building. Thus, we do not consider the placement of the regenerators. Instead, we consider the placement of the buffers because the packet losses occur because of collision without buffers.

III. DATA CENTER NETWORKS WITH HYBRID OPTOELECTRONIC ROUTERS

A. Hybrid optoelectronic routers

In this paper, we discuss an energy saving approach based on the hybrid optoelectronic router proposed by Ibrahim *et al.* [12]. Figure 1 shows a hybrid optoelectronic router. A hybrid optoelectronic router has optical ports and electronic ports, similar to the architecture proposed by Pan *et al.* [25], where

an edge router is attached to the optical packet switch. Using the electronic ports, server racks that send electronic packets can be connected to the hybrid optoelectronic routers without any additional devices. Electrical packets from server racks are buffered in the electrical buffer. They are then converted to optical packets through the optical converter and relayed as an optical packet, which is then relayed between hybrid optoelectronic routers. In addition, optical packets destined for the server racks are stored in the electrical buffer after conversion into electrical packets. The electrical packets are then relayed to the server rack.

In addition, a hybrid optoelectronic router uses the buffer if collision occurs. In the hybrid optoelectronic router, the label processor selects the output port of each packet based on the label attached to it. If no collision occurs, the optical signals are then relayed to the destination port. However, if collision occurs, the optical packets are stored in the electronic buffer after conversion into an electronic packet. The packet is then retransmitted after conversion back into an optical packet.

The hybrid optoelectronic router has two advantages. (1) Unless packet collision occurs, the router can quickly relay the packet without conversion into electronic signals. This leads to low latency and energy consumption. (2) Even if collision occurs, the router can retransmit the packet. As a result, networks with hybrid optoelectronic routers do not require packet transmission to be scheduled to avoid collision. Therefore, networks with hybrid optoelectronic routers are scalable.

B. Data center networks with hybrid optoelectronic routers

Ibrahim *et al.* [12] used hybrid optoelectronic routers to construct a data center network. Figure 2 shows a data center network using hybrid optoelectronic routers. Hybrid optoelectronic routers provide large bandwidth and low latency communication between their ports. Therefore, we use the routers to construct a core network within a data center. In this network, similar to traditional data centers, each server rack has a switch called the top-of-rack (ToR) switch that is connected to all the servers in the rack. Each server rack is connected to the core network by connecting the ToR switch to the hybrid optoelectronic routers. To efficiently use the large bandwidth of the hybrid optoelectronic routers, each router is connected to multiple ToR switches and aggregates the traffic from them. Each ToR switch is also connected to multiple routers to maintain connectivity, even when some routers fail.

In this network, routes can be controlled in a centralized manner based on SDN principles [26]; a central controller collects traffic information and sets the routing tables of the hybrid optoelectronic routers using a protocol based on OpenFlow. A central server periodically collects the traffic information from the network devices within a data center and calculates the routes within a data center network. The central controller then configures the routing tables of hybrid optoelectronic routers. After the routing tables have been configured, the packets are relayed via the configured routes.

The message size of the communication between a central server and the routers is 100 Bytes per flow entry. In the

data center network constructed of the hybrid optoelectronic routers, flows from a server rack can be aggregated into the flows between server racks. Assuming that a data center has 20,000 servers and each server rack contains 80 servers, the total message size needed to configure the routes between all server rack pairs is 6.25 MBytes. This message has only a small impact on the network because its size is very small compared to the capacity of the links. In addition, Benson *et al.* demonstrated that OpenFlow can be used to configure the routes within a data center at 1-s intervals [27].

At the same time it configures the routing tables, the central controller also configures the placement of the buffer; the central controller sends a signal to power on or off the buffer to each hybrid optoelectronic router. An optoelectronic router powers on the buffer immediately after receiving the signal to do so. The time required to power on the buffer is considered to be sufficiently small compared with the time interval to control the routes (e.g., one second), because a recent study demonstrated that even a line card can wake up within 127 ms [28].

In the method proposed in this paper, we reduce energy consumption by centrally controlling these routes; the central controller periodically configures the routes so as to minimize the energy consumption under the constraint that all traffic can be accommodated.

IV. OVERVIEW OF NETWORK ENERGY SAVING USING HYBRID OPTOELECTRONIC ROUTERS

A hybrid optoelectronic router can cut power to its electronic buffer without affecting optical packet switching functionality. Therefore, we consider the energy consumption saved by shutting down the electronic buffers.

The buffer in the optoelectronic router is required only in the following two cases.

- Packets from/to the server rack exist. The electronic packets from the server racks are stored in the electronic buffer and converted into optical packets. Similarly, the optical packets sent to the server rack are converted into electronic packets and stored in the electronic buffer before they are relayed to the server racks.
- Packet collision occurs. In this case, the packets are stored in the electronic buffer and then retransmitted. Packet collision occurs only when packets from the different input ports to the same output ports arrive simultaneously.

Otherwise, we can shut down the electronic buffers without degrading communication performance.

Router buffers can be shut down depending on the traffic routes; because each ToR switch is connected to multiple optoelectronic routers, we have multiple candidate routers that can send and receive packets to/from the server rack. We can also avoid collisions without buffers by setting routes such that at each router, the packets from different ports do not share the same output port. However, the routes needed to avoid packet collision may increase the number of routers that must be powered on. Therefore, we should select the routes according to the energy consumption of the routers themselves as well as their buffers.

On the other hand, the routes within a data center must also provide sufficiently high communication performance; a lack of bandwidth or large delay may degrade the performance of the data center. The communication performance achieved by the routes depends on the traffic demands; even if we set suitable routes for the monitored traffic demands, the routes no longer provide sufficiently high communication performance if the traffic demands change. Thus, we should dynamically reconfigure the routes with respect to changes in traffic.

Therefore, in this paper, the central controller periodically collects traffic information from the network and calculates and configures routes to minimize the energy consumption while providing sufficient bandwidth for all flows. In this approach, the central controller configures the routes for all server rack pairs, including the flows for which the traffic rate is zero. By configuring routes for such flows in advance, even if a new server rack pair starts communication after the routes have been configured, this flow can be accommodated without reconfiguration.

In this approach, we do not need to control the packet transmission schedule because the buffers mitigate packet loss in case of collision. However, because the traffic pattern in a data center can change within a few seconds [27], central routing control must be able to adjust faster than this in order to follow traffic pattern changes. Thus, we need a method to calculate routes immediately. This method is discussed in Section V.

V. ROUTING CALCULATION

In our method, we periodically obtain traffic information and calculate traffic routes. In this section, we discuss a method to calculate suitable routes in each control period.

A. Objective

The objective of our routing is to minimize energy consumption while keeping the communication performance high. In this paper, we use two metrics for communication performance: number of hops and link utilization. That is, we avoid routes where the maximum number of hops exceeds the predefined acceptable number H^{\max} or the utilization of a link becomes larger than the predefined acceptable link utilization U^{\max} .

Our routing method calculates the routes for all source and destination pairs. We define $R_{s,d,l}$ to be a binary variable indicating whether the flow from s to d passes link l ; $R_{s,d,l}$ is 1 if the flow from s to d passes link l , and is 0 otherwise. At the same time, our routing method obtains the powered-on routers and buffers. We define P_n^r to be a binary variable indicating whether router n is powered on or not; P_n^r is 1 if router n is powered on, and otherwise 0. Similarly, P_n^{buf} indicates whether the buffer of router n is powered on or not. In addition, we also define binary variables that indicate whether two flows share the link l or not. We define $P_{s_1,d_1,s_2,d_2,l}^{\text{share}}$ to be a binary variable indicating whether both of the flows from s_1 to d_1 and from s_2 to d_2 pass link l ; $P_{s_1,d_1,s_2,d_2,l}^{\text{share}}$ is 1 if both of the flows from s_1 to d_1 and from s_2 to d_2 pass link l , and is 0 otherwise.

We denote the set of links as L and the set of hybrid optoelectronic routers as N . We also denote the set of server racks as S . We denote the set of outgoing links of node $v \in N \cup S$ as L_v^{out} , and the set of incoming links of node $v \in N \cup S$ as L_v^{in} . The capacity of link l is denoted as C_l . The amount of traffic from server rack s to server rack d is $T_{s,d}$. The energy consumption of a hybrid optoelectronic router without a buffer is E^r and the energy consumption that is additionally required when the buffer is powered on is E^{buf} .

The objective of our routing is formalized as the integer linear programming (ILP) problem,

$$\text{minimize } \sum_{n \in N} (E^r P_n^r + E^{\text{buf}} P_n^{\text{buf}}),$$

with the following constraints:

- The hybrid optoelectronic routers in the path of a flow should be powered on

$$\forall n \in N: P_n^r \geq \frac{1}{|S|^2 |L|} \sum_{s,d \in S, l \in (L_n^{\text{out}} \cup L_n^{\text{in}})} R_{s,d,l}$$

- The buffers of the hybrid routers should be powered on in the following three cases: (1) the packet from a server rack arrives

$$\forall n \in N: P_n^{\text{buf}} \geq \frac{1}{|S|^2 |L|} \sum_{s,d \in S, l \in (L_s^{\text{out}} \cap L_n^{\text{in}})} R_{s,d,l}$$

(2) the packet to a server rack arrives

$$\forall n \in N: P_n^{\text{buf}} \geq \frac{1}{|S|^2 |L|} \sum_{s,d \in S, l \in (L_d^{\text{in}} \cap L_n^{\text{out}})} R_{s,d,l}$$

or (3) packets from different input ports directed to the same output port

$$\forall n \in N, s_1, s_2, d_1, d_2 \in S, l_1 \in L_n^{\text{out}}, l_2 \in L_n^{\text{in}}:$$

$$P_n^{\text{buf}} \geq R_{s_1,d_1,s_2,d_2,l_1}^{\text{share}} - R_{s_1,d_1,s_2,d_2,l_2}^{\text{share}}$$

- All link utilizations should be less than U^{\max}

$$\forall l \in L: \sum_{s,d \in S} T_{s,d} R_{s,d,l} \leq U^{\max} C_l$$

- The number of hops of the flow should be less than H^{\max} .

$$\forall s, d \in S: \sum_{l \in L} R_{s,d,l} \leq H^{\max}$$

- The route for the flow from s to d is a set of continuous links from s to d .

$$\forall s, d \in S : \sum_{l \in L_s^{\text{out}}} R_{s,d,l} = 1$$

$$\forall s, d \in S, n \in N : \sum_{l \in L_n^{\text{out}}} R_{s,d,l} = \sum_{l \in L_n^{\text{in}}} R_{s,d,l}$$

$$\forall s, d \in S : \sum_{l \in L_d^{\text{in}}} R_{s,d,l} = 1$$

(1)

- $P_{s_1,d_1,s_2,d_2,l}^{\text{share}}$ should be 1 if both of the flows from s_1 to d_1 and from s_2 to d_2 pass link l .

$$\forall s_1, s_2, d_1, d_2 \in S, l \in L:$$

$$R_{s_1, d_1, s_2, d_2, l}^{\text{share}} \geq R_{s_1, d_1, l} + R_{s_2, d_2, l} - 2$$

By solving the above problem, we can obtain suitable routes that minimize the energy consumption. However, it is difficult to solve the above problem immediately, especially for a large data center network, because it includes $O(|S|^4|L|)$ binary variables. Therefore, the rest of this section proposes a heuristic method to calculate suitable routes.

B. Basic idea of heuristic route calculation

Instead of calculating the routes of all flows, we calculate the routes of the flows one-by-one. When calculating the route of a flow, we consider the additional energy consumption required, which is calculated as the sum of the energy consumption of the additional powered-on routers and buffers required. By calculating the routes with the minimum additional energy consumption, we avoid routes with large energy consumption.

In addition to the additional energy consumption needed to accommodate the flow, we should consider another point: whether the additional powered-on routers or buffers could be used by other flows. If we calculate the routes so that the newly powered-on routers are likely to be used by other flows, these flows could be accommodated without additional routers, which further reduces energy consumption.

In this paper, to evaluate the probability that the router or buffer could be used by other flows, we define a new metric called the *reuse easiness*. The reuse easiness e_n of hybrid optoelectronic router n is defined by

$$e_n = \sum_{s,d} \frac{N_{s,n,d}^{\text{route}} T_{s,d}}{N_{s,d}^{\text{route}}} \quad (2)$$

where $N_{s,d}^{\text{route}}$ is the number of shortest routes between $s-d$ and $N_{s,n,d}^{\text{route}}$ is the number of routes using node n within the shortest routes between $s-d$.

Reuse easiness indicates the expected amount of traffic passing the hybrid optoelectronic router in the case of the shortest path. If a hybrid optoelectronic router with a small reuse easiness is powered on, it may be used by only a small amount of traffic. On the other hand, a hybrid optoelectronic router with a large reuse easiness accommodates more traffic if it is powered on. Therefore, from the viewpoint of the efficiency of the powered-on hybrid optoelectronic routers, we should select routes that use the hybrid optoelectronic routers with large reuse easiness.

In the proposed method, we consider both the additional energy consumption and reuse easiness. When calculating the route of a flow, we use the following metric instead of the additional energy consumption for route r :

$$W_r^{\text{route}} = \sum_{i=2}^{|r|} W_{l_{i-1}^r, l_i^r} \quad (3)$$

where l_i^r is the i th link of route r and $W_{l_{i-1}^r, l_i^r}$ is defined by

$$W_{l_{i-1}^r, l_i^r} = \begin{cases} E_{l_{i-1}^r, l_i^r}^{\text{additional}} - we_{n_{l_{i-1}^r, l_i^r}} & (E_{l_{i-1}^r, l_i^r}^{\text{additional}} > 0) \\ \epsilon & (\text{otherwise}) \end{cases} \quad (4)$$

where $n_{l_{i-1}^r, l_i^r}$ is the hybrid optoelectronic router connected to links l_{i-1}^r and l_i^r , ϵ is a sufficiently small value, and w is a weight parameter indicating the reuse easiness.

Here, $E_{l_{i-1}^r, l_i^r}^{\text{additional}}$ is the additional energy consumption required to use links l_{i-1}^r and l_i^r , defined as follows.

- If $n_{l_{i-1}^r, l_i^r}$ and its buffer are already powered on, $E_{l_{i-1}^r, l_i^r}^{\text{additional}} = 0$.
- If $n_{l_{i-1}^r, l_i^r}$ is not powered on and both links l_{i-1}^r and l_i^r are links from/to other hybrid optoelectronic routers, $E_{l_{i-1}^r, l_i^r}^{\text{additional}} = E^r$.
- If $n_{l_{i-1}^r, l_i^r}$ is not powered on and either link l_{i-1}^r or l_i^r is a link from/to a server rack, $E_{l_{i-1}^r, l_i^r}^{\text{additional}} = E^r + E^{\text{buf}}$.
- If $n_{l_{i-1}^r, l_i^r}$ is already powered on, but its buffer is not powered on, we check whether the buffer is required. The buffer is required in the following three cases: (1) l_{i-1}^r is a link from the server racks, (2) l_i^r is a link to the server racks, and (3) l_i^r is already used by a flow that does not pass l_{i-1}^r . If the buffer is required, $E_{l_{i-1}^r, l_i^r}^{\text{additional}} = E^{\text{buf}}$. Otherwise, $E_{l_{i-1}^r, l_i^r}^{\text{additional}} = 0$.

By subtracting the reuse easiness from the additional energy consumption in Eq.(4), we reduce the cost of powering on routers or buffers whose reuse easiness is large and avoid powering on optoelectronic routers and buffers whose reuse easiness is small. Note that ϵ is used in Eq.(4) instead of 0 to avoid a route with a large number of hops.

Finally, by selecting the route with the minimum W_r^{route} , we consume only a small amount of additional energy.

C. Steps to calculate routes

Our method calculates the routes of the flows between server rack pairs according to the following steps.

First, we sort the flows in descending order by traffic volume multiplied by the number of hops between the source and destination server racks. This value indicates the total resources required to accommodate the flow. We then calculate the routes of the flows one-by-one starting from the flow with the largest value. As the required total resources of the flow increases, it becomes difficult to determine a route to accommodate the flow when the residual resources are small. Thus, we first search for routes for flows whose required total resources are the largest to ensure that suitable routes for these flows can be determined.

In our method, the route of each flow is calculated by constructing a graph where each vertex corresponds to one of the links in the physical network and edges are added between the vertices whose corresponding physical links are connected to/from the same router. We set the weight of the edges based on $W_{l_{i-1}^r, l_i^r}$. The routes are then calculated over the graph; the shortest path on the graph is the path that minimizes W_r^{route} .

In the rest of this section, we explain how to construct the graph and calculate the routes using this graph.

1) *Constructing the route calculation graph*: In our method, we construct a directed graph whose vertices correspond to links in the physical network. We add an edge from a vertex corresponding to link l_a to a vertex corresponding to a link l_b if the starting point of l_b is the ending point of l_a .

We set the weight of this edge to W_{l_a, l_b} . We also add nodes corresponding to the source and destination server racks and add edges from them to the vertices corresponding to the links connected to them. Using this graph, we can obtain the route with the minimum W_r^{route} by calculating the shortest path from the source to the destination server racks.

The constraint of link utilization is easily considered when constructing this graph; for the route from s to d , the vertex corresponding to link l is eliminated if $T_{s,d} + \sum_{(a,b) \in F^{\text{accommodated}}} T_{a,b} R_{a,b,l} > U^{\text{max}} C_l$, where $F^{\text{accommodated}}$ is the set of flows whose routes have been already decided.

2) *Route calculation*: We can obtain the route with minimum W_r^{route} by calculating the shortest path on the constructed graph. However, the number of hops of the obtained route may be larger than H^{max} . In this case, we must find another route for which the hops do not exceed H^{max} .

In this paper, we use the k-shortest path approach to calculate the route. In this method, we maintain a list of the *uncompleted routes* whose first nodes are the source node of the route to be calculated, but which do not yet include the destination node. The route is calculated by the following steps.

- Step 1 Calculate the sum of edge weights on the shortest path from each node to the destination node using Dijkstra's algorithm. We denote the calculated sum of weights from node n as W_n^{ToDest} .
- Step 2 Calculate the minimum number of hops from each node to the destination node using Dijkstra's algorithm. We define H_n^{ToDest} to be the calculated number of hops from node n .
- Step 3 Add an uncompleted route consisting only of the source node to the list.
- Step 4 Select the uncompleted route r from the list with the smallest $W_r^{\text{UncompletedRoute}} + W_{n_r^{\text{last}}}^{\text{ToDest}}$, where $W_r^{\text{UncompletedRoute}}$ is the sum of weights of r and n_r^{last} is the last node of r .
- Step 5 Add all uncompleted routes generated by adding a node to the r selected in Step 4, and delete r from the list.
- Step 6 Eliminate uncompleted routes whose $H_r^{\text{UncompletedRoute}} + H_{n_r^{\text{last}}}^{\text{ToDest}}$ exceeds H^{max} , where $H_r^{\text{UncompletedRoute}}$ is the number of hops of r .
- Step 7 Check whether there is an entry whose last node is the destination node. If there is, designate that entry as the route from the source to the destination. Otherwise, go back to Step 4.

D. Computational complexity

In our method, the routes are calculated over the route calculation graph, which is constructed of $|L|$ nodes and $|N|p^2$ links where p is the number of ports of each optoelectronic router.

In Steps 1, and 2 of the steps in Section V-C2, the Dijkstra's algorithm is used to calculate the sum of weights or the number of hops from each node to the destination. The

complexity of Dijkstra's algorithm is $O(|N|p^2 + |L|\log|L|)$. The number of hops calculated in Step 2 does not change unless the network topology changes. Thus, the calculation of the number of hops in Step 2 is required only once. On the other hand, the calculation of the sum of weights is required every time, because the weights depend on the powered-on hybrid optoelectronic routers and buffers.

From Steps 4–7, we maintain a list of uncompleted routes. The size of this list is much less than $|L|$. Steps 4–7 are repeated h times where h is the number of hops from the source to the destination. Thus, Steps 4–7 take $O(h|L|)$ time.

In the evaluation in Section VI, a route for each flow in a data center network of 100 hybrid optoelectronic routers and 25 groups of server racks was calculated within 0.5 ms, and the sum of the time required to calculate the routes between all pairs of groups of server racks was 306 ms by a computer with a 2.70 GHz Intel Xeon Processor (E5-2697).

VI. EVALUATION

In this section, we present the results of an evaluation of our method through simulation.

A. Evaluation scenario

1) Evaluation environment:

a) *Network topology*: In our evaluation, we used a torus network according to the previous work about the data center networks constructed of hybrid optoelectronic routers [12]. We used a network constructed of 100 hybrid optoelectronic routers with 16 optical ports and 25 groups of server racks, unless otherwise stated. The 100 hybrid optoelectronic routers construct a core network of 10×10 torus topology, so that each hybrid optoelectronic router pair in the grid topology is connected by four optical links, as shown in Figure 3. Each server rack is connected to four hybrid optoelectronic routers, and all the server racks in the same group are connected to the same hybrid optoelectronic routers. We set the bandwidth of the links between hybrid optoelectronic routers to 100 Gbps.

b) *Energy consumption model*: A hybrid optoelectronic router is constructed of optical label processors, an optical switch, CMOS memory, and optical-to-electronic or electronic-to-optical converters, as shown in Figure 1. When turning on a router without the buffer, the optical label processors and optical switch must be powered on, while the other components can be left shut down as shown in Table I. On the other hand, when the buffer is required, all components should be powered on.

Instead of the energy consumption of each device, it is the ratio of the buffer and converter energy consumption to the total energy consumption of a hybrid optoelectronic router that impacts the performance of our method. Thus, we set the ratio of the energy consumption of the buffer and converters to 0.6 unless otherwise stated, because the buffer and converters are considered to consume more energy than the optical label processor and the optical switch that process packets without conversion between optical and electronic packets.

In addition, we also varied this energy consumption ratio to demonstrate that our method works efficiently, even if the

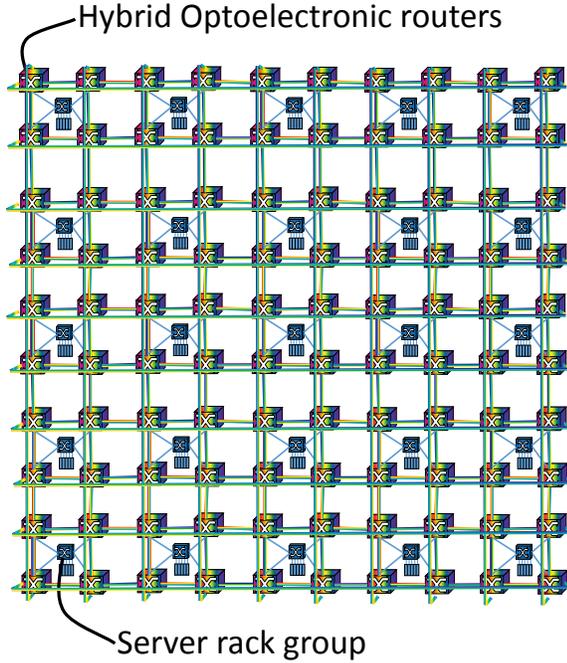


Fig. 3. Network topology used in our evaluation

TABLE I
DEVICES REQUIRED WHEN BUFFERS ARE OFF

Function	Buffer off	Buffer on
Label processors	Required	Required
Optical switch	Required	Required
CMOS memory	Not required	Required
Converters	Not required	Required

buffer and converters consume less energy than the optical label processors and optical switch.

c) Traffic: In this evaluation, we generated traffic only between selected server rack pairs, because each server communicates with only a limited number of servers concurrently [29]. To generate traffic, we first selected a predefined number of server rack group pairs. We then generated the traffic rates of these pairs by generating uniform random numbers. Finally, we scaled the generated traffic rate so that the total traffic rate matched the predefined value.

We set the number of communicating server rack group pairs to 300, which is half of all the server rack group pairs unless otherwise stated. In addition, we evaluated the impact of the number of communicating server rack group pairs by changing the number from 100 to 600.

The total traffic rate was set considering that servers tend to communicate with the servers in the same rack; in the cloud data centers, intra-rack communication occupies about 80 % of total traffic [30]. We set the total traffic rate to 4 Tbit/s unless otherwise stated, assuming that the data center has 20,000 servers, each server has a 1 Gbit/s port and the traffic rate between server racks is 20 % of the capacities of the server ports. We also evaluated the impact of the traffic rate by changing the total traffic rate from 1 to 8 Tbit/s.

In this evaluation, 10 traffic patterns were generated for each

combination of pair numbers and total traffic rates. We show the averages, maximums, and minimums of the 10 results. In the figures in this section, the maximum and minimum values are shown by error bars.

d) Parameters: In the proposed method, we use two parameters to indicate the sufficiency of communication performance: the acceptable number of hops H^{\max} and acceptable link utilization U^{\max} . In this evaluation, we set H^{\max} and U^{\max} to sufficiently large values to investigate the potential reduction of the energy consumption achieved by our method. Hence, H^{\max} was set to 27 and U^{\max} to 1.0. In addition, we also evaluated the impact of H^{\max} by changing H^{\max} from 7 to 13. We did not evaluate the impact of U^{\max} , because it is similar to the impact of the traffic rate.

Our method also has parameter w , which is the weight of reuse easiness. In this evaluation, we set w to a sufficiently small value in order to select routes with the smallest additional energy. By setting w to a sufficiently small value, we only compare reuse easiness when routes with the same additional energy exist.

2) Compared methods: In this study, we compared our method with the following methods.

a) Method to minimize the number of powered-on routers: In this method, we minimize the number of powered-on hybrid optoelectronic routers without considering the energy consumption of the buffers. Unnecessary optoelectronic routers and buffers are then shut down. Because minimizing the number of powered-on routers takes a long time to calculate, we calculated the routes using the same approach as our method described in Section V; the flow routes are calculated one-by-one using the graph constructed as described in Section V-C. In contrast to our method, when setting a weight for each edge as in Eq. (4), we used an additional energy defined by

$$E_{l_{i-1}^r, l_i^r}^{\text{additional}} = \begin{cases} 0 & n_{l_{i-1}^r, l_i^r} \text{ is powered on} \\ E^r + E^{\text{buf}} & \text{Otherwise} \end{cases}$$

b) Method to use the shortest paths: In this method, we calculated the flow routes one-by-one, similar to our method, but when calculating a route for a flow, we do not consider the energy consumption; the route is calculated using the shortest paths over the physical network after eliminating the links whose utilizations become larger than U^{\max} . By comparing our method with this method, we can clarify the impact of considering the energy consumption.

3) Metrics: In this evaluation, we compared the energy consumption of the core network achieved by each method. We normalized the energy consumption by the consumption of all hybrid optoelectronic routers and buffers together.

B. Comparison with the optimal solution

First, we compare the results of our heuristic method with the optimal solution obtained by ILP described in Section V-A. For this comparison, we used the small network topology shown in Figure 4, which is constructed of 16 hybrid optoelectronic routers, and four groups of server racks, because it is impractical to obtain the optimal solution for the network

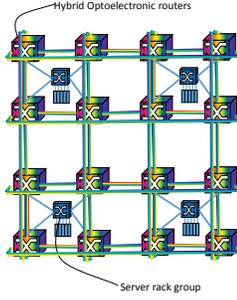


Fig. 4. Network topology used in our evaluation

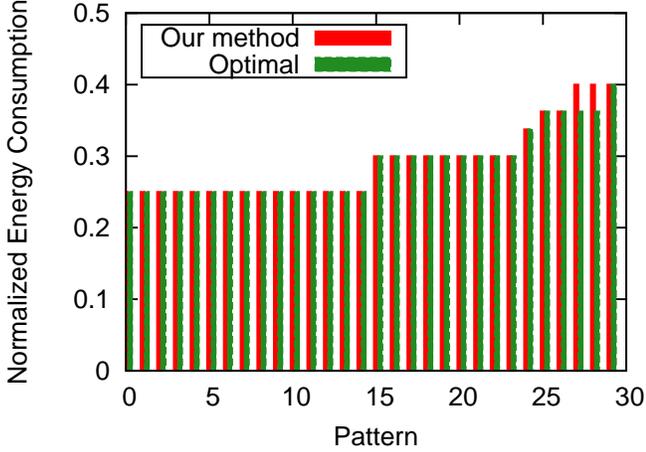


Fig. 5. Comparison of our heuristic method with ILP

topology shown in Figure 3. We set the bandwidth of the links between hybrid optoelectronic routers to 100 Gbps. Each hybrid optoelectronic router pair was connected by one link.

We generated the traffic between all server rack groups randomly so that the total traffic rate was 120, 240, or 360 Gbps. For each total traffic rate value, we generated 10 traffic patterns. Figure 5 compares the energy consumption achieved by our heuristic and the optimal solution. In this figure, the vertical axis indicates the energy consumption normalized by the consumption of all hybrid optoelectronic routers and buffers together. We plot the results for all 30 cases of the generated traffic. Figure 5 shows that our heuristic method achieves the same energy consumption as the optimal solution in most cases. Even when the energy consumption achieved by our heuristic method is larger than the optimal solution, the difference is only the energy consumption of one buffer. That is, this result demonstrates that our heuristic method obtains nearly optimal solutions.

C. Comparison of the energy consumption

We compare the energy consumption achieved by each method. Figure 6 shows the normalized energy consumption of the optical and electronic devices that were achieved by each method. In this graph, the vertical axis is the average of the normalized energy consumption for all 10 traffic patterns. Furthermore, “OurMethod” indicates the results for our method,

“MinRouter” indicates the result for the method that minimizes the powered-on routers, “MinHop” indicates the result for the method that uses the shortest paths, and “ALL-ON” indicates the energy consumption when all hybrid optoelectronic routers and buffers are powered on. From this figure, it is clear that our method can significantly reduce energy consumption below that of the other methods. In particular, the energy consumption of the electronic devices such as CMOS memory is reduced significantly by our method. This is a result of the reduction of the number of powered-on buffers. The energy consumption of the optical devices cannot be reduced much, unlike the energy consumption of the electronic devices. This is because a certain number of routers are required to keep the connectivity between all server rack group pairs, and the number of powered-on routers cannot be reduced as much as the number of powered-on buffers.

Figure 7 shows the number of powered-on devices for each method. This figure clarifies that our method significantly reduces the number of powered-on buffers, while the number of powered-on routers equals that of “MinRouter”. This is because our method configures routes to avoid collisions without powering on new buffers. On the other hand, “MinRouter” does not consider the powered-on buffers, and hence most of the buffers on the powered-on routers are powered-on.

Table II shows the properties of the routes calculated by each method in more detail. In this table, the number of used physical links indicates the number of optical links passed by at least one flow between server racks. The average or maximum link utilization indicates the average or maximum link utilization of the used links, respectively. Similarly, the average or maximum number of hops respectively indicates the average or maximum number of hops between all server rack pairs. For each metric, Table II shows the average, maximum, and minimum values over all traffic patterns generated in this evaluation.

This table indicates that “MinHop” achieves the smallest number of hops, but uses the most physical links. The average link utilization of “MinHop” is the smallest. These results indicate that “MinHop” cannot use links efficiently. Compared with “MinHop”, our method and “MinRouter” achieve much higher link utilization. That is, our method and “MinRouter” efficiently use the links to reduce the powered-on devices.

This table also indicates that the number of hops is significantly greater for our method and “MinRouter” than it is for “MinHop”. Even if the number of hops for a route is greater than the shortest route, our method and “MinRouter” use that route if it can achieve the smallest energy consumption by reducing the number of powered-on routers. However, a large number of hops causes a high latency or high bit error rate, and may be avoided. In our method, a large number of hops can be avoided by setting H^{\max} to a small value. The impact of H^{\max} on the energy saving is discussed in Section VI-G.

D. Impact of total traffic volume

We investigated the impact of total traffic volume on the energy consumption achieved by our method. Figure 8 shows the energy consumption achieved by each method when the

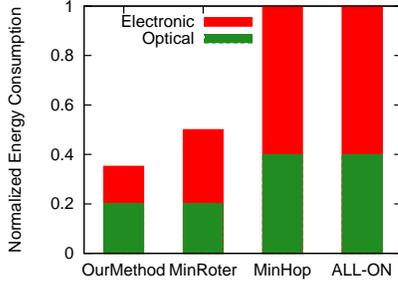


Fig. 6. Energy consumption achieved by each method

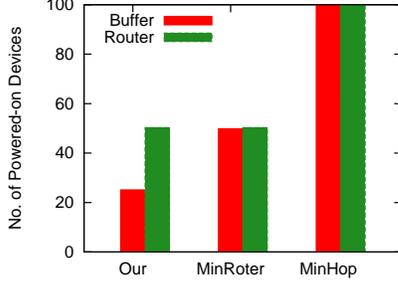


Fig. 7. Number of devices necessary for each method

TABLE II
PROPERTIES OF THE ROUTES

		OurMethod	MinRouter	MinHop
Number of used physical links	Average	373.9	355.8	402.2
	Min	343	333	396
	Max	409	377	407
Average link utilization	Average	0.85	0.82	0.33
	Min	0.84	0.79	0.32
	Max	0.86	0.84	0.34
Maximum link utilization	Average	1.00	1.00	1.00
	Min	1.00	1.00	0.99
	Max	1.00	1.00	1.00
Average number of hops	Average	8.18	7.57	3.33
	Min	7.92	7.17	3.33
	Max	8.63	7.80	3.33
Maximum number of hops	Average	21.9	22.5	7
	Min	17	21	7
	Max	29	25	7

total traffic rate is set from 1 to 8 Tbit/s. In this figure, the vertical axis is the normalized energy consumption achieved by each method, and the horizontal axis is the total traffic rate. This figure shows that the energy consumption becomes large as the total traffic rate increases. This is because more hybrid optoelectronic routers are required to accommodate a larger amount of traffic. Comparing the energy consumption achieved by each method, our method consumes only less than a half the energy of “ALL-ON,” even when the total traffic rate is 8 Tbit/s, while the normalized energy consumption of “MinRouter” becomes about 70% that of “ALL-ON.” Thus, our method saves energy even when the traffic rate becomes large.

To discuss the reason why our method can reduce the energy consumption even for large amounts of total traffic, we investigated the number of powered-on devices. Figure 9

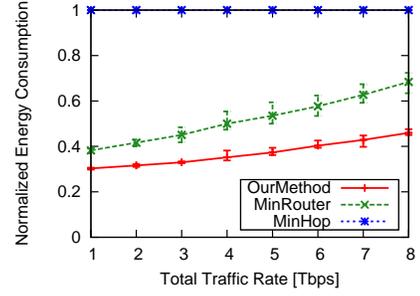
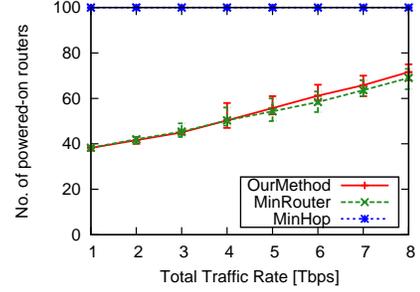
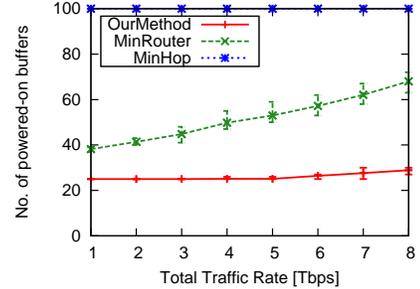


Fig. 8. Impact of total traffic rate on energy consumption



(a) Hybrid optoelectronic routers



(b) Buffers

Fig. 9. Impacts of total traffic rate on the number of powered-on devices

shows the number of powered-on routers or buffers in each method. In this figure, the vertical axis is the number of powered-on hybrid optoelectronic routers or buffers, and the horizontal axis is the total traffic rate. This figure shows that the number of powered-on hybrid optoelectronic routers or buffers increases for both of our method and “MinRouter”. It also shows that the rate of increase in the number of powered-on buffers in our method is significantly slower than “MinRouter,” and the number of powered-on buffers in our method is always the smallest among all methods. On the other hand, the increase of the number of powered-on routers in our method is faster than it is in “MinRouter.” This is because our method avoids powering on buffers if we can achieve lower energy consumption by avoiding it, even if some additional hybrid optoelectronic routers are powered on. For example, if a route that requires one additional powered-on hybrid optoelectronic router can avoid the collision of the packets on the hybrid optoelectronic routers whose buffers are off, our method selects that route, because powering on new buffers consumes more energy than powering on one hybrid

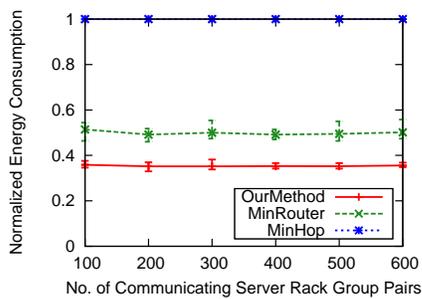


Fig. 10. Impact of the number of communicating server rack group pairs on energy consumption

optoelectronic router. As a result, our method uses the smallest number of buffers, even though more routers are powered on. On the other hand, the other methods power on buffers on almost all of the powered-on routers. As a result, our method can achieve significantly smaller energy consumption than the other methods.

E. Impact of number of communicating server rack group pairs

We also changed the number of communicating server rack group pairs to investigate its impact on energy consumption. Figure 10 shows the result. In this figure, the vertical axis is the normalized energy consumption and the horizontal axis is the number of communicating server rack group pairs. From this figure, no clear impact of the number of pairs on energy consumption is shown. This is because we first allocate routes, even for flows between server rack group pairs that are not communicating at the time the traffic is monitored. We then power on the routers and buffers required by such flows. By allocating routes for such flows, we can accommodate the traffic of these flows if the server rack group pairs start communicating with each other before the routes are reconfigured. Therefore, we conclude that the number of communicating server racks does not have an impact on the number of powered-on devices and energy consumption.

F. Impacts of device energy consumptions

We investigated the energy consumption of devices by changing the ratio of the buffer energy consumption. Figure 11 shows the impact of the buffer energy consumption on the total energy consumption achieved by our method and “MinRouter.” In this figure, the vertical axis is the normalized total energy consumption, and the horizontal axis is the ratio of the buffer energy consumption. This figure demonstrates that the normalized energy consumption achieved by our method becomes large as the buffer energy consumption decreases. This is caused by the fact that the rate of reduction in the energy consumption by shutting down buffers slows. If the ratio of the buffer energy consumption is 0, our method achieves the same energy consumption as “MinRouter.” However, if a buffer consumes a non-negligible amount of energy, our method consumes less energy than “MinRouter,” even if the energy consumption of a buffer is smaller than the optical

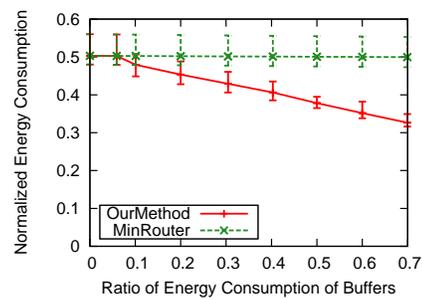
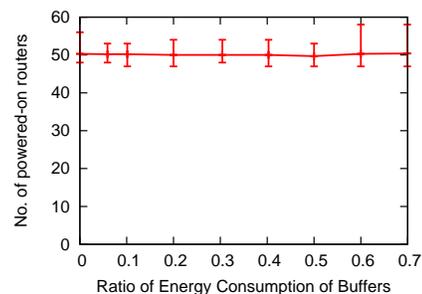
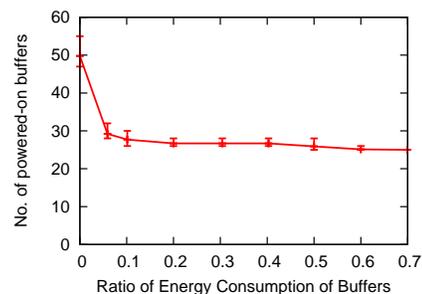


Fig. 11. Impact of buffer energy consumption on the energy consumption achieved by our method



(a) Hybrid optoelectronic routers



(b) Buffers

Fig. 12. Impact of buffer energy consumption on the number of powered-on devices

devices. Even if the ratio of the buffer energy consumption is 0.3, our method achieves a normalized energy consumption of 0.45, while “MinRouter” cannot achieve a normalized energy consumption that is less than 0.5.

We also investigated the impact of the buffer energy consumption on the route calculated by our method. Figure 12 shows the impact of the buffer energy consumption on the number of powered-on routers and buffers. This figure indicates that buffer energy consumption does not have a significant impact on the number of powered-on routers and buffers unless the energy consumption of the buffers is 0. The number of powered-on buffers is about 50 when the energy consumption of the buffers is 0, but that reduces to about 25 even if the ratio of the energy consumption of the buffer is 0.05. This is because our method set routes so as to reduce the number of powered-on buffers unless the buffer consumes energy.

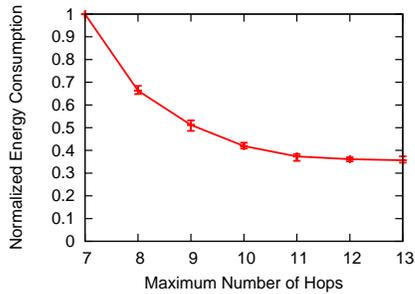


Fig. 13. Impact of the maximum number of hops constraint on the energy consumption of our method

G. Impact of the number of hops constraint

We investigated the impact of the maximum number of hops constraint. Figure 13 shows the impact of maximum number of hops on the total energy consumption achieved by our method. In this figure, the vertical axis is the normalized total energy consumption, and the horizontal axis is the maximum number of hops. The maximum number of hops achieved by “MinHop” is seven as shown in Table II. Thus, the maximum number of hops cannot be set to the value less than seven.

This figure indicates that the energy consumption can be reduced by allowing a large number of hops. If the maximum number of hops is set to the same value achieved by “MinHop,” the normalized energy consumption becomes 1. However, by allowing a slightly larger number of hops, the normalized energy consumption is significantly reduced. Even when the maximum number of hops is set to 10, the normalized energy consumption becomes about 0.40. That is, our method reduces energy consumption without a large number of hops.

H. Impact of considering the reuse easiness

Finally, we investigated the impact of considering reuse easiness. Figure 14 compares the energy consumed by our method with and without considering reuse easiness. Our method without considering reuse easiness calculated routes by setting w to 0. In this comparison, we varied the total traffic rate from 1 to 8 Tbit/s. Figure 14 indicates that reuse easiness has only a small impact on the energy consumed. This is because all hybrid optoelectronic routers have a similar reuse easiness because all server rack groups generate a similar rate of traffic, even though the traffic rate of each flow is set to a random value.

Therefore, we also investigated the impact of reuse easiness in a case where the hybrid optoelectronic routers had different reuse easiness. To do this, we investigated a case where only some of the server racks are active because server racks may be powered off to save energy consumption when the load on the data center is small. Figure 15 shows the impact of the reuse easiness when five randomly selected server rack groups are active, where it is assumed that unused servers are shut down but the servers are not migrated because of the large overhead. In this case, the traffic is generated only between the active server rack group pairs. Unlike the case where

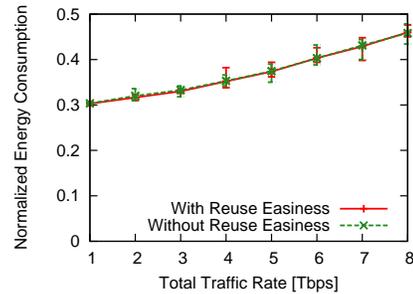


Fig. 14. Impact of considering the reuse easiness (All server racks are active)

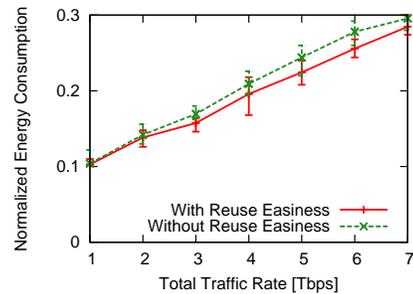


Fig. 15. Impact of considering the reuse easiness (5 server rack groups are active)

all server rack groups are active, the hybrid optoelectronic routers have different reuse easiness; the reuse easiness of the hybrid optoelectronic routers on the shortest path between the active server rack groups increases, while that of the other routers decreases. As shown in Figure 16, we generated traffic between all five active server rack pairs and changed the total traffic rate from 1 to 7 Tbit/s.

Figure 15 indicates that our method considering reuse easiness consumes less energy when five server racks are active. This is because our method considers the reuse easiness values on the routers or buffers on the shortest paths between the active server racks. As a result, the powered-on routers and buffers are used by the other paths and hence effectively used.

Therefore, we conclude that reuse easiness does not impact energy consumption when all servers generate similar rates of traffic and all hybrid optoelectronic routers have similar reuse easiness. On the other hand, when hybrid optoelectronic routers have different reuse easiness (e.g., when only a small number of server racks are active), considering reuse easiness reduces energy consumption.

VII. CONCLUSION

In this paper, we proposed a heuristic route method that calculates traffic routes to reduce the energy consumption of a data center network constructed of hybrid optoelectronic routers. In this method, we accommodate traffic within the data center network with a small number of routers and buffers and shut down unnecessary routers and devices. Through simulation, we demonstrated that our method reduces the energy consumption significantly compared to a method that

does not consider buffer energy consumption as well as one that simply calculates the shortest traffic routes through the network.

One of our future research topics is to evaluate the effectiveness of our method on various data center network environments, such as other network topologies.

ACKNOWLEDGMENTS

The authors would like to thank all the people in the program “Basic technologies for high-performance optoelectronic hybrid packet router,” and particularly Ryo Takahashi from NTT Corporation.

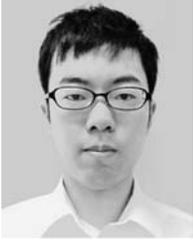
REFERENCES

- [1] A. Greenberg, J. Hamilton, D. A. Maltz, and P. Patel, “The cost of a cloud: research problems in data center networks,” *ACM SIGCOMM Computer Communication Review*, vol. 39, pp. 68–73, Jan. 2009.
- [2] D. Abts, M. Marty, P. Wells, P. Klausler, and H. Liu, “Energy proportional datacenter networks,” in *Proceedings of the 37th annual international symposium on computer architecture (ISCA 2010)*, pp. 338–347, June 2010.
- [3] S. J. B. Yoo, “Optical packet and burst switching technologies for the future photonic internet,” *Journal of Lightwave Technology*, vol. 24, pp. 4468–4492, Dec. 2006.
- [4] N. Farrington, G. Porter, S. Radhakrishnan, H. H. Bazzaz, V. Subramanya, Y. Fainman, G. Papen, and A. Vahdat, “Helios: a hybrid electrical/optical switch architecture for modular data centers,” *ACM SIGCOMM Computer Communication Review*, vol. 40, pp. 339–350, Oct. 2010.
- [5] C. Guillemot, M. Renaud, P. Gambini, C. Janz, I. Andonovic, R. Bauknecht, B. Bostica, M. Burzio, F. Callegati, M. Casoni, D. Chiaroni, F. Clerot, S. L. Danielsen, F. Dorgeuille, A. Dupas, A. Franzen, P. B. Hansen, D. K. Hunter, A. Kloch, R. Krahenbuhl, B. Lavigne, A. L. Corre, C. Raffaelli, M. Schilling, J.-C. Simon, and L. Zucchelli, “Transparent optical packet switching: The european ACTS KEOPS project approach,” *Journal of Lightwave Technology*, vol. 16, pp. 2117–2134, Dec. 1998.
- [6] Z. Zhu, V. J. Hernandez, M. Y. Jeon, J. Cao, Z. Pan, and S. J. B. Yoo, “Rf photonics signal processing in subcarrier multiplexed optical-label switching communication systems,” *Journal of Lightwave Technology*, vol. 21, pp. 3155–3166, Dec. 2003.
- [7] Z. Pan, H. Yang, Z. Zhu, and S. J. B. Yoo, “Demonstration of an optical-label switching router with multicast and contention resolution at mixed data rates,” *IEEE Photonics Technology Letters*, vol. 18, pp. 307–309, Jan. 2006.
- [8] H. J. Chao and K. Xi, “Bufferless optical cros switches for data centers,” in *Proceedings of Optical Fiber Communication Conference (OFC 2011)*, pp. 1–3, Mar. 2011.
- [9] O. Alparslan, S. Arakawa, and M. Murata, “Comparison of packet switch architectures and pacing algorithms for very small optical RAM,” *International Journal on Advances in Internet Technology*, pp. 159–169, June 2010.
- [10] R. Takahashi, T. Nakahara, K. Takahata, H. Takenouchi, T. Yasui, N. Kondo, and H. Suzuki, “Ultrafast optoelectronic packet processing for asynchronous, optical-packet-switched networks,” *Journal of Optical Networking*, vol. 3, pp. 914–930, Dec. 2004.
- [11] H. Takenouchi, R. Urata, T. Nakahara, T. Segawa, H. Ishikawa, and R. Takahashi, “First demonstration of a prototype hybrid optoelectronic router,” in *proceedings of the 35th European Conference on Optical Communication (ECOC 2009)*, pp. 1–2, Sept. 2009.
- [12] S. A. Ibrahim, H. Ishikawa, T. Segawa, T. Nakahara, Y. Suzaki, and R. Takahashi, “100-Gb/s optical packet switching technologies for data center networks,” in *proceedings of Photonics in Switching, PW4B.2*, July 2014.
- [13] M. Chen, H. Jin, Y. Wen, and V. C. M. Leung, “Enabling technologies for future data center networking: a primer,” *IEEE Network*, pp. 8–15, Aug. 2013.
- [14] L. Wang, F. Zhang, J. A. Aroca, A. V. Vasilakos, K. Zheng, C. Hou, D. Li, and Z. Liu, “GreenDCN: a general framework for achieving energy efficiency in data center networks,” *IEEE journal on Selected Areas in Communications*, pp. 4–15, June 2014.
- [15] D. Kilazovich, P. Bouvry, and S. U. Khan, “DENS: data center energy-efficient network-aware scheduling,” *Cluster Computing*, pp. 65–75, Mar. 2013.
- [16] M. Zhang, C. Yi, B. Liu, and B. Zhang, “GreenTE: Power-aware traffic engineering,” in *Proceedings of the 18th IEEE International Conference on Network Protocols (ICNP 2010)*, pp. 21–30, Oct. 2010.
- [17] B. Heller, S. Seetharaman, P. Mahadevan, Y. Yiakoumis, P. Sharma, S. Banerjee, and N. Mckeown, “ElasticTree: Saving energy in data center networks,” in *Proceedings of the 7th USENIX conference on Networked systems design and implementation (NSDI 2010)*, pp. 17–32, Apr. 2010.
- [18] F. Vismara, V. Grkovic, F. Musumeci, M. Tornatore, and S. Bregni, “On the energy efficiency of IP-over-WDM networks,” in *Proceedings of IEEE Latin-American Conference on Communications (LATINCOM 2010)*, pp. 1–6, Sept. 2010.
- [19] G. Shen and R. Tucker, “Energy-minimized design for IP over WDM networks,” *IEEE/OSA Journal of Optical Communications and Networking*, vol. 1, pp. 176–186, June 2009.
- [20] F. Idzikowski, S. Orłowski, C. Raack, H. Woesner, and A. Wolisz, “Saving energy in IP-over-WDM networks by switching off line cards in low-demand scenarios,” in *Proceedings of the 14th Conference on Optical Network Design and Modeling (ONDM 2010)*, pp. 1–6, Feb. 2010.
- [21] E. Yetginer and G. Rouskas, “Power efficient traffic grooming in optical WDM networks,” in *Proceedings of Global Telecommunications (GLOBECOM 2009)*, pp. 1–6, Nov. 2009.
- [22] A. Nag and M. Tornatore, “Transparent vs. translucent optical network design with mixed line rates,” in *Proceedings of OFC, OW17*, Mar. 2009.
- [23] Z. Zhu, “Mixed placement of 1R/2R/3R regenerators in translucent optical networks to achieve green and cost-effective design,” *IEEE Communications Letters*, vol. 15, pp. 752–754, May 2011.
- [24] Z. Zhu, X. Chen, F. Ji, L. Zhang, F. Farahmand, and J. P. Jue, “Energy-efficient translucent optical transport networks with mixed regenerator placement,” *Journal of Lightwave Technology*, vol. 30, pp. 3147–3156, Oct. 2012.
- [25] Z. Pan, H. Yang, J. Yang, J. Hu, Z. Zhu, J. Cao, K. Okamoto, S. Yamano, V. Akella, and S. J. B. Yoo, “Advanced optical-label routing system supporting multicast, optical ttl, and multimedia applications,” *Journal of Lightwave Technology*, vol. 23, pp. 3270–3280, Oct. 2005.
- [26] K. Kitayama, S. Debnath, Y. Yoshida, R. Takahashi, A. Hiramatsu, Y. Ohsita, M. Murata, and M. Hayashitani, “Green and smart data center networks based on optical packet switching with flow control,” *IEEE Summer Topicals*, July 2013.
- [27] T. Benson, A. Anand, A. Akella, and M. Zhang, “MicroTE: Fine Grained Traffic Engineering for Data Centers,” in *Proceedings of ACM CoNEXT*, pp. 1–12, Dec. 2011.
- [28] T. Pan, T. Zhang, J. Shi, Y. Li, L. Jin, F. Li, J. Yang, B. Zhang, and B. Liu, “Towards zero-time wakeup of line cards in power-aware routers,” in *Proceedings of IEEE INFOCOM*, pp. 190–198, May 2014.
- [29] C. Kachris and I. Tomkos, “A survey on optical interconnects for data centers,” *IEEE Communications Surveys & Tutorials*, vol. 14, pp. 1021–1036, Nov. 2012.
- [30] T. Benson, A. Akella, and D. A. Maltz, “Network traffic characteristics of data centers in the wild,” in *Proceedings of Internet Measurement Conference*, pp. 267–280, Nov. 2010.



Yuichi Ohsita received his M.E. and Ph.D. degrees in information science and technology from Osaka University in 2005 and 2008.

He is currently an Assistant Professor in the Graduate School of Information Science and Technology, Osaka University. His research interests include traffic engineering, traffic prediction, and data center networks. He is a Member of IEEE, the Association for Computing Machinery (ACM), and IEICE.



Takamichi Nishijima received his M.E. and Ph.D. degrees in information science and technology from Osaka University in 2011 and 2014. He is a Member of IEEE and IEICE.



Yuki Koizumi received his M.E. and Ph.D. degrees in information science and technology from Osaka University, Japan, in 2006 and 2009, respectively. He is currently an Assistant Professor at the Graduate School of Information Science and Technology, Osaka University, Japan. His research interests include traffic engineering in photonic networks and biologically inspired networking. He is a Member of IEEE and IEICE.



Masayuki Murata received his M.E. and D.E. degrees in information and computer sciences from Osaka University in 1984 and 1988, respectively.

In April 1984, he joined the Tokyo Research Laboratory of IBM Japan as a Researcher. From September 1987 to January 1989, he was an Assistant Professor with the Computation Center, Osaka University. In February 1989, he moved to the Department of Information and Computer Sciences, Faculty of Engineering Science, Osaka University. From 1992 to 1999, he was an Associate Professor with the

Graduate School of Engineering Science, Osaka University, and since April 1999, he has been a Professor. He moved to the Graduate School of Information Science and Technology, Osaka University in April 2004. He has published more than 300 papers in international and domestic journals and conferences. His research interests include computer communication networks, performance modeling, and evaluation.

He is a Fellow of IEICE, and a Member of IEEE, ACM, the Internet Society, and IPSJ.