JOURNAL OF
Optical Communications
and Networking

Optical circuit switched three-stage twisted-folded Clos-network design model guaranteeing admissible blocking probability

Ryotaro Taniguchi,[1] Takeru Inoue,[2] Kazuya Anazawa,[1,2] and Eiji Oki[1,*]

[1]Graduate School of Informatics, Kyoto University, Kyoto 606-8501, Japan
[2]NTT Network Innovation Laboratories, NTT Corporation, Kanagawa, Japan
*oki@i.kyoto-ac.jp

Received 9 July 2024; revised 23 September 2024; accepted 29 September 2024; published 21 October 2024

Some data center networks have already started to use optical circuit switching (OCS) with potential performance benefits, including high capacity, low latency, and energy efficiency. This paper addresses a switching network design to maximize the network radix, i.e., the number of terminals connected to the network under the condition that a specified number of identical switches with the size $N \times N$ and the maximum admissible blocking probability are given. Previous work presented a two-stage twisted and folded Clos network (TF-Clos) with a blocking probability guarantee for OCS, which has a larger network radix than TF-Clos with a strict-sense non-blocking condition. Expanding the number of stages allows for enhancing the network radix. This paper proposes a model designing an OCS three-stage TF-Clos structure with a blocking probability guarantee to increase the network radix compared to the two-stage TF-Clos. We formulate the problem of obtaining the network configuration that maximizes the network radix as an optimization problem. We conduct an algorithm based on an exhaustive search to obtain a feasible solution satisfying the constraints of the optimization problem. This algorithm identifies the structure with the largest network radix in non-increasing order to avoid unnecessary searches. Numerical results show that the proposed model achieves a larger network radix than the two-stage model.

https://doi.org/10.1364/JOCN.535282

1. INTRODUCTION

Data center networks consist of hardware components, such as switches and routers, that enable operators to provide services to users. Some data center networks have already deployed optical circuit switching (OCS) with potential performance benefits [1–4]. OCS enables high-capacity and low-latency data communication. In addition, OCS is energy efficient, which can contribute to lower energy consumption. Terminals, such as top-of-racks (ToRs) or aggregation switches, can be directly connected via OCS in data center networks.

A Clos network can be used in a data center network [5], a multi-stage structure designed to provide a highly scalable, non-blocking connection. In a traditional Clos network, the architecture consists of three stages: input layer, intermediate layer, and output layer. A request is routed through switches located in each layer without affecting existing optical connections. Throughout this paper, we do not consider rearranging existing optical connections, as we focus on maintaining communication stability. Clos network designs with a large network radix, i.e., switching network size, will be required to accommodate future traffic growth. The network radix is defined as the number of terminals, each with a transmitter and receiver pair connected to the network.

There are several types of Clos networks. All the Clos variants considered in this paper are fully optical, with all switches being OCS; each switch is itself strict-sense non-blocking (SNB). We assume that we design a switching network under the given number of identical switches with the size $N \times N$. The work in [6] presented a folded Clos network (F-Clos). Figure 1 shows the structure of two-stage F-Clos, which combines the input and output stages of a traditional three-stage Clos network. Folding these stages reduces the overall number of stages, simplifying the architecture and decreasing the switching network's complexity. The work in [7] studied three-stage F-Clos, which replaces the input-output layer in two-stage F-Clos with another smaller two-stage F-Clos. A book [8] introduced the variants of F-Clos. In Fig. 1, we consider that each optical switch has two sets of ports; each pair of ports belonging to different sets can be connected via the optical switch. Each set of ports is called a *side*. The input-output layer has k switches, each connecting n transmitters and n receivers. The intermediate layer has m switches. In F-Clos,

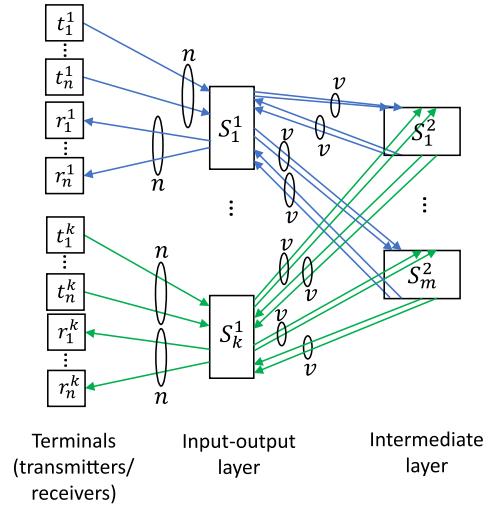

Fig. 1. Structure of two-stage F-Clos. The input-output layer has k switches, each connecting n transmitters and n receivers. The intermediate layer has m switches. One side of each switch in the input-output layer is connected to n transmitters and n receivers. The other side of each switch in the input-output layer connects one side of each switch in the intermediate layer with v links.

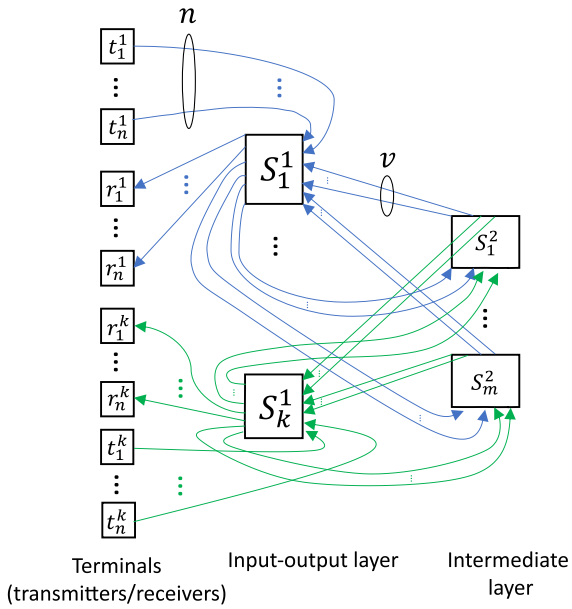

Fig. 2. Structure of two-stage TF-Clos. The input-output layer has k switches, each connecting n transmitters and n receivers. The intermediate layer has m switches. One side of each switch in the input-output layer connects one side of each switch in the intermediate layer with v links.

one side of each switch in the input-output layer is connected to n transmitters and n receivers. The other side is connected to switches in the intermediate layer with v links. A switch in the input-output layer requires $2n$ ports on one side and vm ports on the other side. A switch in the intermediate layer requires vk ports on each side.

The works in [9,10] presented a two-stage twisted and folded Clos network (TF-Clos), which is a more flexible switching network structure than two-stage F-Clos. Figure 2 presents the structure of TF-Clos. We use the same definition of *side* of a switch as described in F-Clos. In TF-Clos, both sides of each switch in the input-output layer are connected to terminals; one side functions as either a set of input or output ports, unlike F-Clos. TF-Clos inherits the characteristic of F-Clos, where the input and output stages of a traditional three-stage Clos network are combined into a single layer. Two-stage F-Clos and two-stage TF-Clos (2TF) are all designed with SNB conditions [8,10]. 2TF has a larger network radix than the two-stage F-Clos [9,10] for the same number of switches. This is because the constraints of the switch ports used to design 2TF are more relaxed than those of two-stage F-Clos, while the feasible region of n, k, v, and m for 2TF depicted in Fig. 2, each of which expresses a feasible structure of 2TF, covers that of two-stage F-Clos. When we have a targeted network radix of a switching network, 2TF can require fewer switches than two-stage F-Clos [11]. For example, to achieve the network radix of 390 with the SNB condition, 2TF requires 55 switches with $n = 13$, $k = 30$, $v = 1$, and $m = 25$ in Fig. 2, whereas two-stage F-Clos requires 58 switches with $n = 10$, $k = 39$, $v = 1$, and $m = 19$ in Fig. 1; the required number of switches in 2TF is 0.95 times fewer than that of two-stage F-Clos. Section 3 provides a detailed description of the 2TF in Fig. 2.

The blocking probability is a performance indicator that expresses the percentage of blocked requests in a network. A request may not reach its destination in a switching network since no internal link is available on the route. The work in [12] approximately analyzed the blocking probability in a traditional Clos network but did not give a theoretical upper bound of the blocking probability.

As the network radix, i.e., the number of terminals connected to the network, increases, more internal switches are required to prevent internal blocking and maintain connection quality. Therefore, addressing the design challenges of increasing the network radix while maintaining an admissible blocking probability is essential. An admissible blocking probability denotes the maximum tolerable probability that a terminal request is blocked, which is given by a network operator or designer.

The works in [13–15] introduced 2TF with a blocking probability guarantee. They aim to maximize the network radix of 2TF under the given number of identical switches with the size $N \times N$. A network designer prepares a set of available identical switches without the need to prepare a variety of different-sized switches. This paper adopts this specific assumption, which instills confidence in its practicality. The work in [13] provided a one-step, 2TF design model that satisfies an admissible blocking probability. In this model, the acceptance of a connection request is determined by checking the utilization of links from the input-output layer to the intermediate layer without rearranging existing connections in the network. This decision is made once for each connection request, so the model is referred to as a one-step, 2TF design model. Under the given number of identical switches, the network radix of the one-step, 2TF model with a blocking probability guarantee is larger than that of 2TF with the SNB condition. From another perspective, if we relax the SNB condition, we need fewer switches to obtain a given network

radix. The work in [14] provided a two-step, 2TF design model with a blocking probability guarantee by dividing the set of intermediate-layer switches into two groups, where a request admission is judged by checking the utilization of links from the input-output layer to the intermediate layer in two steps. Considering the admissible blocking probability in each group, the two-step, 2TF model with a blocking probability guarantee has a larger network radix than the one-step, 2TF model with a blocking probability guarantee. The work in [15] provided an S-step ($S \geq 1$), 2TF design model that generalizes the number of steps with a blocking probability guarantee. By generalizing the number of steps, the S-step, 2TF model with a blocking probability guarantee has a larger network radix than the model with at most two steps. The design models [13,14] restricted the number of stages to two. No study addressed a TF-Clos design model considering more than two stages.

Questions arise. *Can we provide a design model of TF-Clos to increase the network radix, guaranteeing an admissible blocking probability by relaxing the number of stages to more than two under the given number of switches? How much network radix can we increase in the design model of TF-Clos with more than two stages compared with that of two stages?*

This paper proposes a model to design the three-stage TF-Clos (3TF) for OCS that guarantees an admissible blocking probability to maximize the network radix. Our design model uses identical switches with the size $N \times N$, each with N input ports and N output ports, which is the same assumption as that of [13–15]. 3TF replaces the input-output layer in 2TF with another smaller 2TF. 3TF regards a set of first- and second-stage switches as a giant switch in the input-output layer. The number of ports of such a giant switch can increase compared with that of an $N \times N$ switch. The input-output layer and the intermediate layer in the proposed model are connected in the same way that the input-output layer and intermediate layer in 2TF are connected. We formulate the proposed design model as an optimization problem. We derive the theoretical upper bound of blocking probability for designing 3TF. To solve the problem, we employ an exhaustive search-based algorithm. This algorithm identifies the structure with the largest network radix in non-increasing order instead of finding a feasible solution for every combination of natural number decision variables while adhering to specified constraints, including the guarantee of blocking probability. Thus, once we find a feasible solution during the search, we ensure that it is optimal and the algorithm terminates. We shorten the process by appropriately ordering the decision variables to be searched. This paper is the first study in which we design 3TF with a blocking probability guarantee by deriving the theoretical upper bound of blocking probability. Numerical results show that the proposed model for designing 3TF has a larger network radix than the baseline model for designing 2TF under the SNB condition or with a guarantee of maximum blocking probability. Increasing the number of stages in TF-Clos from two to three increases the computation time to design TF-Clos. The OCS network constructed by the proposed 3TF design model applies to a data center network that directly connects terminals, such as ToRs or aggregation switches, while maintaining blocking quality.

Table 1. TF-Clos Design Models

Number of Stages	SNB	Blocking Probability Guarantee
Two stages	[9,10]	[13–15]
Three stages		This work

Table 1 categorizes TF-Clos design models. The works in [9,10] designed 2TF under the SNB condition. The works in [13–15] designed 2TF with a blocking probability guarantee. This work contributes to providing for designing 3TF under the blocking probability guarantee and the SNB condition.

The rest of this paper is organized as follows. Section 2 describes the related works. Section 3 presents existing design models of 2TF. Section 4 introduces our proposed design model of 3TF with guaranteeing admissible blocking probability. Section 5 presents numerical results. Section 6 describes the applicability of the proposed model. Finally, we conclude this paper in Section 7.

2. RELATED WORKS

Previous works have studied an unfolded Clos network (UF-Clos) as the basis of a Clos network [5,7,8,16,17]. UF-Clos is a traditional Clos network consisting of three stages. We call this type of a Clos network UF-Clos to distinguish it from F-Clos. The work in [16] introduced round-robin-based dispatching schemes for UF-Clos packet switches to achieve high throughput. The work in [17] improved the schemes in [16] to save on consumed hardware resources. Unlike F-Clos, each transmitter and receiver pair in a UF-Clos are connected to different switches. Our work considers F-Clos applicable to data center networks; F-Clos allows a transmitter and receiver pair to be connected to a common switch.

Current data center networks widely employ packet-switching F-Clos [18–21]. The work in [18] investigated load balancing in F-Clos. The work in [19] introduced a rebalancing algorithm in F-Clos. The work in [20] presented relative definitions of the system reliability of F-Clos. The work in [21] introduced networks randomly connecting certain switch levels. Our work addresses OCS, which differs from packet switching, in F-Clos for data center networks.

With the advancement of optical technology, high bandwidth, low latency, and reduced energy consumption for the data center network are expected [22]. OCS can achieve lower latency compared to electrical packet switches. The work in [23] introduced an optical network with high scalability and low power consumption while providing sufficient bandwidth capacity for data center networks. The work in [24] analyzed an effective topology reconfiguration strategy called OCBridge, which can use link resources efficiently in data center networks. The work in [25] presented an optical switching architecture based on micro electro mechanical systems (MEMS), Flexnet, with low latency, high throughput performance, and high flexibility for data center networks. This paper attempts to design a switching network using optical switches for future data center networks.

In circuit switching, when a connection is established between an endpoint pair, the links used by the established

connection are exclusively occupied. Unlike packet switching, these links cannot be used for communication between other endpoint pairs. In [26], the study analyzed the blocking probability in a circuit-switched network. Additionally, [23] presented a data center network based on OCS. Our study addresses a switching network handling OCS, designing TF-Clos while ensuring an admissible blocking probability.

OCS networks find diverse applications. The work in [27] introduced a methodology for designing a large-scale non-blocking OCS network employing an F-Clos architecture within a data center network tailored for cloud computing. Additionally, the works in [26,28] presented network construction for artificial intelligence (AI) systems using optical switches and robot arms. A direct-connect fabric requires a network that accommodates thousands of terminals. Our work can be applied to networks for cloud computing in a data center and AI systems utilizing OCS.

The blocking probability is a significant performance indicator related to efficiency and scalability. The work in [29] studied the SNB operation of a three-stage Clos network. In [10], the study presented 2TF under the SNB condition. In contrast to our work, these studies focus on the SNB condition, where the blocking probability equals zero.

Several studies [30–35] have conducted investigations to assess the blocking probability in Clos networks. The study in [30] delved into the blocking probability within switching networks employing a reservation algorithm. Similarly, [31] examined the blocking probability in switching networks utilizing Erlang fixed point approximation. The studies in [32,33] addressed analyzing the blocking probability in Clos-network packet switching networks employing virtual output queues. The work in [34] assessed the blocking probability employing artificial neural networks, while [35] developed non-blocking conditions and probabilistic analytical models for blocking within a wavelength-selective switched-based Clos network. Despite their comprehensive analyses of the blocking probability in high-performance Clos networks, these studies did not establish a theoretical bound on the blocking probability.

The network radix of the network can be increased by relaxing the SNB condition while ensuring an admissible blocking probability. The work in [13] introduced a TF-Clos design model with a blocking probability guarantee in a one-step manner. The work in [14] introduced a two-step TF-Clos design model with a blocking probability guarantee, where a request admission is judged by checking the utilization of links from the input-output layer to the intermediate layer in two steps. The work in [15] introduced a TF-Clos design model that generalizes the number of steps with a blocking probability guarantee. The works in [13–15] represent 2TF. Our study extends the number of stages from two to three.

3. EXISTING DESIGN MODEL OF A TWO-STAGE TWISTED-FOLDED CLOS NETWORK

A. Two-Stage Twisted-Folded Clos Network

Figure 2 shows a two-stage TF-Clos for OCS consisting of input-output and intermediate layers [9,10]. 2TF consists of $k+m$ switches, each of which has N input ports and N output ports. In the input-output layer, there are k switches denoted by S_i^1, $i \in [1, k]$. A switch in the input-output layer connects n transmitters denoted by t_j^i, $i \in [1, k]$, $j \in [1, n]$, and n receivers denoted by r_j^i, $i \in [1, k]$, $j \in [1, n]$. In the intermediate layer, there are m switches denoted by S_i^2, $i \in [1, m]$. One side of each switch in the input-output layer connects one side of each switch in the intermediate layer with v links. The SNB condition is given by

$$2\left\lfloor \frac{n-1}{v} \right\rfloor + 1 \le m. \tag{1}$$

Equation (1) is the necessary and sufficient condition for SNB; the proof was given by [5,8]. It indicates that the condition is given by n, v, and m, but does not involve k.

The work in [9,10] defines a design model of TF-Clos with the SNB condition, which determines n, k, m, and v to maximize the network radix of nk, where the number of $N \times N$ switches, a, is given, where $k+m \le a$. The network radix of 2TF is equivalent to nk. When we maximize nk, $k+m$ does not always equal a since the constraints of the number of available ports of switches in the input-output layer and the intermediate layer must be satisfied, as described in Section 3.B. In this case, $a-(k+m)$ $N \times N$ switches are not used.

B. 2TF Design Model with Blocking Probability Guarantee

There are two models [13,14] for OCS, which we name one-step and two-step design models, to determine the structure of 2TF that guarantees an admissible blocking probability. The one-step and two-step models ensure an admissible blocking probability in one and two steps, respectively. We describe the one-step design model for 2TF [13], which is the most basic one to clarify the essential features of changing from 2TF to 3TF. In this section, we call the one-step design model for 2TF a 2TF design model.

The 2TF design model [13] determines decision variables n, n^{snb}, k, m, and v to construct 2TF so that it can maximize the network radix of nk while guaranteeing an admissible blocking probability, ϵ. The 2TF design model also considers Fig. 2, whereas it newly introduces a decision variable of n^{snb}, compared to the TF-Clos design model with the SNB condition [9,10]. $n^{\mathrm{SNB}}(\le n)$ is the maximum number of available terminals in each switch of the input-output layer to guarantee the admissible blocking probability, at most n^{SNB} connections outgoing from each switch of the input-output layer to the intermediate layer. If more than n^{SNB} connection requests arrive, the excessive requests are blocked. The number of $N \times N$ switches, a, is given. It handles blocking one time. We assume that a request generated by a transmitter connected to a terminal is active with probability p and inactive with probability $1-p$. This request generation process follows an independent and identically distributed (i.i.d.) pattern. There are no assumptions about each request's destination; the destination is arbitrary. ϵ denotes the admissible blocking probability. N, a, p, and ϵ are given parameters. The 2TF design problem is given by

$$\max \quad nk, \tag{2a}$$

$$\text{subject to} \qquad n + vm \le N, \tag{2b}$$

$$vk \le N, \tag{2c}$$

$$\sum_{w=n^{\text{snb}}+1}^{n} \binom{n}{w} p^w (1-p)^{n-w} \le \epsilon, \tag{2d}$$

$$2\left\lfloor \frac{n^{\text{snb}}-1}{v} \right\rfloor + 1 \le m, \tag{2e}$$

$$k + m \le a, \tag{2f}$$

$$n, n^{\text{snb}}, k, v, m \in \mathbb{N}. \tag{2g}$$

Equation (2a) is the objective function to maximize the network radix of 2TF. Equations (2b) and (2c) represent the constraints of the number of available ports of switches in the input-output layer and the intermediate layer, respectively. Equation (2d) indicates that the blocking probability does not exceed ϵ. The left side of Eq. (2d) is derived according to a binomial distribution that the probability that more than n^{snb} out of n requests are active. The proof that this indicates the upper bound of the blocking probability for 2TF is in [13]. Equation (2e) corresponds to the SNB condition. In Eq. (2e), n^{snb} replaces n presented in Eq. (1). Equation (2f) indicates that the number of switches used in the network must not exceed a. Equation (2g) represents the type of decision variable.

4. PROPOSED DESIGN MODEL WITH BLOCKING PROBABILITY GUARANTEE

This section presents our proposed model for designing 3TF with a blocking probability guarantee to increase the network radix compared to 2TF for OCS.

A. Overview of Three-Stage Twisted and Folded Clos Network

Figure 3 shows 3TF consisting of input-output and intermediate layers, which we design in the proposed model. Table 2 shows the notations used in the proposed model. 3TF comprises $k(k'+m')+m$ switches, each of which has N input ports and N output ports. The input-output layer corresponds to the first and second stages, and the intermediate layer corresponds to the third stage. The input-output layer comprises k switching-network components, each of which has $k'+m'$ switches, and the intermediate layer includes m switches. We call this an input-output-layer switching-network component (IOC). IOCs and the third-stage switches are connected in the same way as TF-Clos, and so are the first- and second-stage switches. $nk'k$ transmitters and $nk'k$ receivers are connected to all the first-stage switches. The network radix of 3TF is equivalent to $nk'k$. $S^1_{i_1 i_2}$ denotes the i_2th first-stage switch of i_1th IOC, where $i_1 \in [1, k]$ and $i_2 \in [1, k']$. $S^2_{i_1 i_2}$ denotes the i_2th second-stage switch of the i_1th IOC, where $i_1 \in [1, k]$ and $i_2 \in [1, m']$. S^3_i denotes the ith third-stage switch, where $i \in [1, m]$. $t^{i_1 i_2}_j$ and $r^{i_1 i_2}_j$ denote the jth transmitter and the jth receiver, respectively, where $j \in [1, n]$, $i_1 \in [1, k']$, and $i_2 \in [1, k]$; $S^1_{i_1 i_2}$ connects n transmitters and n receivers. $S^1_{i_1 i_2}$ is connected to $S^2_{i_1 i'_2}$ with v' outgoing links and v' incoming links, where $i_1 \in [1, k]$, $i_2 \in [1, k']$, and $i'_2 \in [1, m']$. $S^2_{i_1 i_2}$ is connected to S^3_i with v outgoing links and v incoming links, where $i_1 \in [1, k]$, $i_2 \in [1, m']$, and $i \in [1, m]$. $S^2_{i_1 i_2}$ has at most z output ports outgoing to third-stage switches and at most z input ports incoming from third-stage switches, where $i_1 \in [1, k]$ and $i_2 \in [1, m']$. The total number of output ports from each IOC to third-stage switches is at most zm', and the total number of input ports incoming from third-stage switches to each IOC is at most zm'.

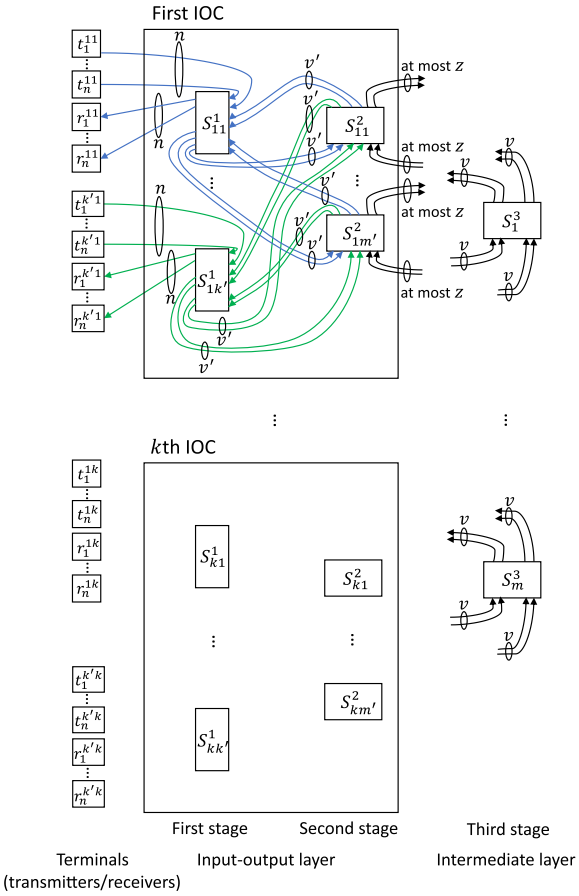

Fig. 3. Structure of 3TF.

We use the same assumption of the request arrival as described in Section 3.B. A request generation process follows an independent and identically distributed (i.i.d.) pattern with the probability of p. No assumptions about each request's destination are imposed.

We describe our proposed design model for 3TF with a two-step blocking probability guarantee. The proposed model also supports the case of the SNB condition ($\epsilon = 0$). We introduce another admissible blocking probability, η, which is a decision variable, where $\eta \le \epsilon \le 1$ in addition to the given parameter of ϵ. We consider two steps to block the requests for connection. A request can be blocked in the first and second

Table 2. List of Notations

Parameters	Descriptions
N	Number of ports in a switch
a	Number of switches that we can use
p	Probability that a request is active
ϵ	Admissible blocking probability in the network
Variables	**Descriptions**
n	Natural number variable that is the number of transmitters or receivers connected to a first-stage switch
k	Natural number variable that is the number of IOCs in the network
k'	Natural number variable that is the number of first-stage switches in an IOC
n_1^{snb}	Natural number variable that is the number of ports that satisfy the SNB condition in an IOC
n_2^{snb}	Natural number variable that is the number of ports that satisfy the SNB condition in the network
m'	Natural number variable that is the number of second-stage switches in an IOC
m	Natural number variable that is the number of third-stage switches in the network
v'	Natural number variable that is the number of links from (to) a first-stage switch outgoing to (incoming from) a second-stage switch, respectively
v	Natural number variable that is the number of links from (to) an IOC outgoing to (incoming from) a third-stage switch, respectively
z	Natural number variable that is the maximum number of output (input) ports at a second-stage switch outgoing to (incoming from) third-stage switches, respectively
η	Real number variable that is the upper bound of the blocking probability in each IOC

stages (first-step) due to the SNB condition restricted by η. If a request is not blocked in the first-step, the request goes to the third stage (second step). In the second step, the request can also be blocked to the SNB condition restricted by ϵ, which is the original admissible blocking probability.

B. Formulation

Given the number of $N \times N$ switches, a, p, and ϵ, we formulate an optimization problem to determine n, n_1^{snb}, n_2^{snb}, k', k, v', v, m', m, $z \in \mathbb{N}$, and $\eta \in \mathbb{R}^+$ as

$$\max \quad nk'k, \tag{3a}$$

$$\text{subject to} \quad n + v'm' \le N, \tag{3b}$$

$$v'k' + z \le N, \tag{3c}$$

$$vk \le N, \tag{3d}$$

$$vm \le zm', \tag{3e}$$

$$\sum_{w=n_1^{\mathrm{snb}}+1}^{n} \binom{n}{w} p^w (1-p)^{n-w} \le \eta, \tag{3f}$$

$$2\left\lfloor \frac{n_1^{\mathrm{snb}} - 1}{v'} \right\rfloor + 1 \le m', \tag{3g}$$

$$1 - (1-\eta) \times \left(1 - \sum_{w=n_2^{\mathrm{snb}}+1}^{nk'} \binom{nk'}{w} p^w (1-p)^{nk'-w}\right) \le \epsilon, \tag{3h}$$

$$2\left\lfloor \frac{n_2^{\mathrm{snb}} - 1}{v} \right\rfloor + 1 \le m, \tag{3i}$$

$$k(k' + m') + m \le a, \tag{3j}$$

$$\eta \le \epsilon \le 1, \tag{3k}$$

$$n, n_1^{\mathrm{snb}}, n_2^{\mathrm{snb}}, k', k, v', v, m', m, z \in \mathbb{N}, \tag{3l}$$

$$\eta \in \mathbb{R}^+, \tag{3m}$$

where $\mathbb{N}$ is a set of natural numbers, and $\mathbb{R}^+$ is a set of non-negative real numbers. Equation (3a) is the objective function to maximize the network radix of 3TF, i.e., to maximize the number of terminals that are connected to the input and output ports of the switching network. Equations (3b)–(3d) represent the constraints of the number of available ports of switches in the first, second, third stages, respectively. Equation (3e) indicates that the total number of output (input) ports outgoing from (incoming to) an IOC to (from) third-stage switches is at least vm, respectively. We can handle Eqs. (3c) and (3e) with the following one integrated constraint:

$$v'k' + \left\lceil \frac{vm}{m'} \right\rceil \le N, \tag{4}$$

where $z = \left\lceil \frac{vm}{m'} \right\rceil$ is determined by v, m, and m'. Equation (3f) indicates that the blocking probability in the first stage does not exceed η. Equation (3g) corresponds to the SNB condition in the first stage. Equation (3h) indicates that the blocking probability in the third stage does not exceed ϵ; the probability that a request is not blocked in the third stage must be larger than $1 - \epsilon$, which is expressed by

$$(1-\eta)\left(1 - \sum_{w=n_2^{\mathrm{snb}}+1}^{nk'} \binom{nk'}{w} p^w (1-p)^{nk'-w}\right) \ge 1 - \epsilon, \tag{5}$$

which derives Eq. (3h). Equation (3i) corresponds to the SNB condition in the third stage. Equation (3j) indicates that the number of switches used in the network must not exceed a. Equation (3k) expresses the range of η, which does not exceed ϵ. Equations (3l) and (3m) represent the types of decision variables. We note that Eqs. (3a)–(3m) also deal with the 3TF design for the SNB condition by setting $\epsilon = 0$.

Let P_2^{B} be the probability of the number of requests exceeding the SNB condition in the second step. P_2^{B} is given by

$$P_2^{\mathrm{B}} = 1 - (1-\eta)\left(1 - \sum_{w=n_2^{\mathrm{snb}}+1}^{nk'} \binom{nk'}{w} p^w (1-p)^{nk'-w}\right). \tag{6}$$

Theorem 1. P_2^{B} indicates the upper bound of the blocking probability for 3TF.

Proof. The upper bound of the blocking probability in the first stage is η. Then, the probability that requests go to the third stage switches is $1-\eta$. If the number of requests that go to the third stage is at most n_2^{snb}, no blocking occurs. P_2^{B} indicates the probability of the number of requests exceeding the SNB condition in the second step. If all $w - n_2^{\mathrm{snb}}$ requests are blocked in the third stage, where $w \in [n_2^{\mathrm{snb}}, nk']$, the blocking probability is P_2^{B}. Thus, P_2^{B} indicates the upper bound of blocking probability for 3TF. □

n_1^{snb} is the maximum number of available terminals in each first-stage switch, whereas n_2^{snb} is the maximum number of available terminals in each IOC. If more than n_1^{SNB} connection requests arrive at a first-stage switch, the excessive requests are blocked. If more than n_2^{SNB} connection requests arrive at an IOC, the excessive requests are blocked. The number of available terminals in one first-stage switch of an IOC can differ from that of another first-stage switch in the same IOC. Since the structure with $n_1^{\mathrm{snb}}k' \geq n_2^{\mathrm{snb}}$ is allowed, not all $n_1^{\mathrm{snb}}k'$ connections can be outgoing from an IOC. From a wiring point of view, the same number of terminals are connected to each first-stage switch in the same IOC.

P_2^{B} is a function of η, which denotes $f(\eta)$.

Theorem 2. $f(\eta)$ is a non-decreasing function for η.

Proof. Let $f'(\eta)$ be the derivative of $f(\eta)$. $f'(\eta)$ is given by

$$f'(\eta) = 1 - \sum_{w=n_2^{\mathrm{snb}}+1}^{nk'} \binom{nk'}{w} p^w (1-p)^{nk'-w}. \tag{7}$$

Equation (7) indicates that $f'(\eta)$ is non-negative. Thus, $f(\eta)$ is a non-decreasing function of η. □

From Theorem 2, we can get an optimal solution by substituting the smallest η satisfying Eq. (3f) into Eqs. (3h) and (3k). Therefore, Eq. (3f) is rewritten as follows:

$$\sum_{w=n_1^{\mathrm{snb}}+1}^{n} \binom{n}{w} p^w (1-p)^{n-w} = \eta. \tag{8}$$

C. Design Algorithm

To solve the optimization problem in Eqs. (3a)–(3m), we conduct an algorithm based on an exhaustive search for the integer decision variables, as shown in Algorithm 1. It exhaustively searches for the structure of the largest network radix in its non-increasing order instead of finding a feasible solution for every combination of the natural number decision variables. If a feasible solution with a larger network radix is found, a feasible solution with a smaller network radix is no longer optimal. Thus, once we find a feasible solution during the search, we ensure that it is optimal and the algorithm terminates. We shorten the process by appropriately ordering the decision variables to be searched.

We keep the maximum value of network radix $nk'k$ in the already searched space by checking each feasible set of the decision variables. To avoid unnecessary searches, Algorithm 1 searches for decision variables n, k', and k in descending order of $nk'k$, starting with the combination having the largest value of $nk'k$ (line 3), and it ends the search when it finds a feasible solution. Moreover, Algorithm 1 can avoid unnecessary searches by determining whether the conditions are satisfied at each point in the *for* loop (lines 5, 7, 9, 11, 13, and 15). If $N \leq a$, $nk'k$ is at most $N(N-1)(N-1)$ without considering any constraints in the proposed 3TF design model, and otherwise $Na(a-1)$. Similarly, each of n_1^{snb}, v, and v' is at most N. n_2^{snb} is at most $N(N-1)$ if $N \leq a$, and $N(a-1)$ otherwise. Each of m and m' is at most $a-1$. When completing the search, we obtain the structure of the proposed model with the maximum network radix. The overall computational time complexity to solve the optimization problem is $O(N^8a^2)$ if $N \leq a$, and $O(N^5a^5)$ otherwise.

5. NUMERICAL RESULTS

This section evaluates the proposed 3TF design model regarding network radix, computation time, and the required number of switches, compared with two design models for 2TF, one-step (1S) and two-step (2S) models, as baseline models. In the 2TF design model with 1S (2TF-1S) [13], a request admission is judged by checking the utilization of links from the input-output layer to the intermediate layer in one step, whereas, in the 2TF design model with 2S (2TF-1S) [14], it is performed in two steps. 2TF-1S and 2TF-2S are applicable for $\epsilon > 0$. Only 2TF-1S is applicable for $\epsilon = 0$; 2TF-2S manages request blocking in a two-step manner only when $\epsilon > 0$.

We use algorithms based on an exhaustive search to obtain the network structure for each model, where the proposed 3TF design model runs Algorithm 1; we code them in

Algorithm 1. Algorithm Based on Exhaustive Search Largest Network Radix in Its Non-increasing Order

Input: N, a, p, ϵ
Output: n, n_1^{snb}, n_2^{snb}, k', k, v', v, m', m
1: nk′k_array ← [[1, 1, 1], [1, 1, 2], … , [$N-1$, $N-1$, N]]
2: Sort in descending order with respect to the value of the multiplication of three values in the array with nk′k_array.
3: **for** nk′k_array **do**
4: **for** v **do**
5: **if** Eq. (3d) is satisfied **then**
6: **for** n_1^{snb} **do**
7: **if** Eqs. (8) and (3k) are satisfied **then**
8: **for** n_2^{snb} **do**
9: **if** Eq. (3h) is satisfied **then**
10: **for** m **do**
11: **if** Eq. (3i) is satisfied **then**
12: **for** m' **do**
13: **if** Eq. (3j) is satisfied **then**
14: **for** v' **do**
15: **if** Eqs. (3b), (3g), and (4) are satisfied **then**
16: **return** n, n_1^{snb}, n_2^{snb}, k', k, v', v, m', m

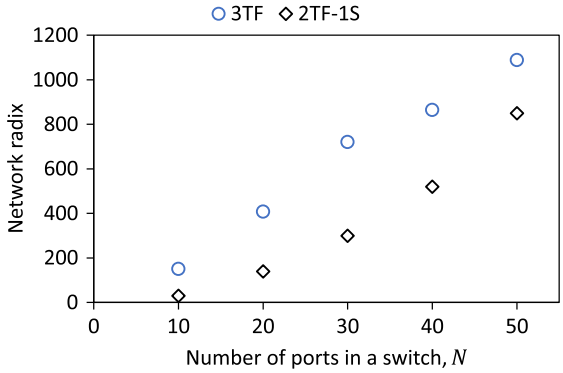

Fig. 4. Dependency of the network radix on N with $a = 200$ and $\epsilon = 0$.

C++. Our used hardware platform is an Intel Core i7-7700 CPU at 3.60 GHz 4-core CPU and 32 GB memory to run Algorithm 1.

A. Network Radix

We consider a scenario where a service provider wants to design a data center network to maximize the network radix while guaranteeing an admissible blocking probability, given the number of $N \times N$ switches, in Sections 5.A and 5.B.

Figure 4 shows the dependency of the network radix on N under the SNB condition. We compare the switching capacities of 3TF and 2TF-1S with $N = 10$, 20, 30, 40, and 50. We set $a = 200$. The network radix of TF-Clos increases by changing the number of stages from two to three.

We investigate the network radix dependent on a under the SNB condition, as shown in Fig. 5. We set $N = 40$. The larger a becomes, the greater the difference in the network radix between two stages and three stages is. Increasing the number of stages from two to three increases the value of a with which the network radix saturates. The appearance of the ceiling for the network radix of 3TF when a becomes greater than 2500 is due to the lack of ports for connecting the new links while satisfying the SNB condition. In 2TF-1S, there is no free port to connect the links when a is greater than 200.

Figure 6 observes the dependency of network radix on ϵ with $N = 40$, $a = 2400$, and $p = 0.5$; it also presents the network radix under the SNB condition as a reference. The network radix of 3TF is larger than those of 2TF-1S and 2TF-2S over

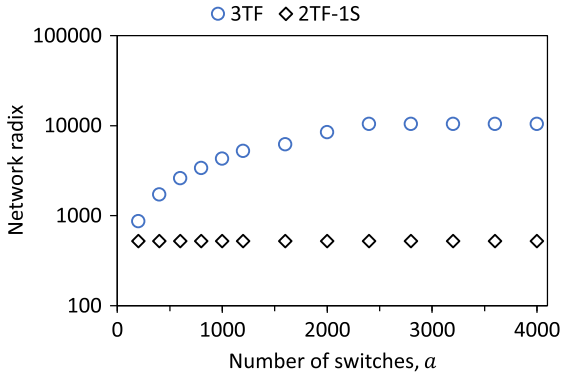

Fig. 5. Dependency of network radix on a with $N = 40$ and $\epsilon = 0$.

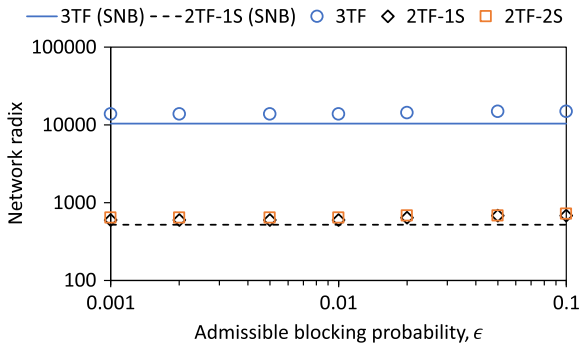

Fig. 6. Dependency of network radix on ϵ with $N = 40$, $a = 2400$, and $p = 0.5$.

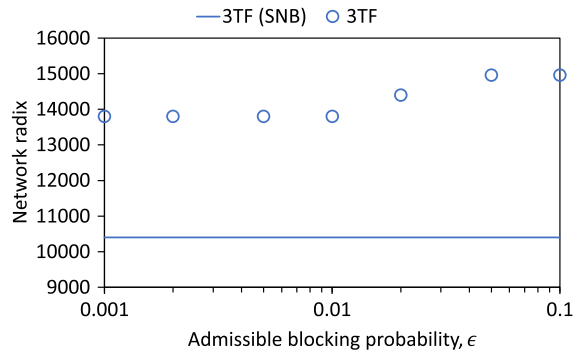

Fig. 7. Dependency of network radix on ϵ with $N = 40$, $a = 2400$, and $p = 0.5$ for 3TF.

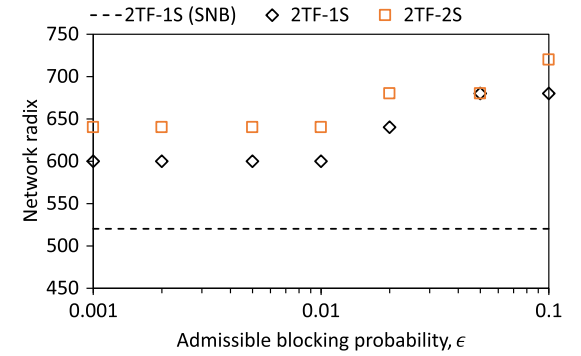

Fig. 8. Dependency of network radix on ϵ with $N = 40$, $a = 2400$, and $p = 0.5$ for 2TF-1S and 2TF-2S.

our examined ϵ. To see the ϵ dependency in detail, we change the scale of the vertical axis of Fig. 6 and depict the results in Figs. 7 and 8. As ϵ increases, the network radix of each model increases. By changing ϵ from zero to a positive value, the network radix increases in 2TF-1S, 2TF-2S, and 3TF. The network radix of 2TF-2S is larger than that of 2TF-1S, but its difference is much smaller than the difference between those of 2TF-1S and 3TF.

We evaluate the network radix along with varying p with $N = 40$, $a = 2400$, and $\epsilon = 0.1$, as shown in Fig. 9. As p increases, the network radix of each model decreases. The network radix of 2TF-2S is slightly larger than or equal to that of 2TF-1S, but that of 3TF is much larger than that of 2TF-1S.

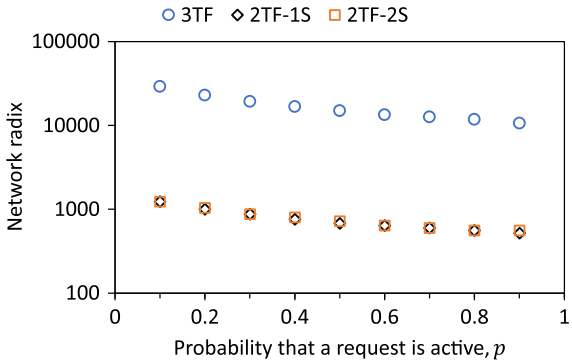

Fig. 9. Dependency of network radix on p with $N = 40$, $a = 2400$, and $\epsilon = 0.1$.

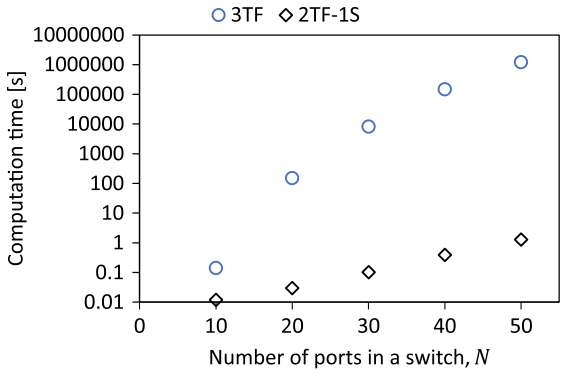

Fig. 10. Dependency of computation time on N with $a = 200$ and $\epsilon = 0$.

B. Computation Time

Figure 10 shows the computation time on N under the SNB condition. The computation time of 3TF increases more sharply than that of 2TF-1S in N. When $N = 50$, it takes about 13 days to solve the 3TF design problem. We note that a service provider seeks to create a data center network that optimizes the network radix and ensures a predetermined blocking probability before its service operation. The network structure design is permitted to take a few weeks.

Figure 11 shows the dependency of computation time on a with $N = 40$ and $\epsilon = 0$. The computation time of designing 3TF is larger than that of 2TF-1S; the number of *for* loops that pass before finding a solution in the algorithm of the 3TF design model is greater than that of the 2TF-1S design model. The larger a becomes, the smaller the computation time of the 3TF design model is. Increasing a promotes relaxing the condition that the total number of switches is less than a; it is easier to satisfy the condition in the conditional branching of the algorithm. The easier it is to satisfy the condition, the sooner the algorithm obtains the solution. The greater the number of *for* loops until the algorithm obtains the solution, the greater the reduction in computation time derived from increasing a. For this reason, we observe decreasing the computation time for the 3TF design model with an increase in a.

We investigate the computation time dependent on ϵ with $N = 40$, $a = 2400$, and $p = 0.5$, as shown in Fig. 12. The computation time of 3TF tends to increase slightly as ϵ increases. The larger ϵ is, the more the number of sets of variables that satisfy Eq. (3h). As a result, the number of *for* loops that pass in the algorithm increases. Conversely, in the 2TF-1S, the computation time decreases as ϵ increases due to the small number of *for* loops. Thus, the computation time is larger when $\epsilon > 0$ than under the SNB condition in the 3TF design model, and the opposite is true in the 2TF-1S design model. The computation time of 2TF-2S also decreases as ϵ increases for the same reason as the 2TF-1S, but the computation time is larger than the 2TF-1S; 2TF-2S handles a larger number of decision variables than 2TF-1S.

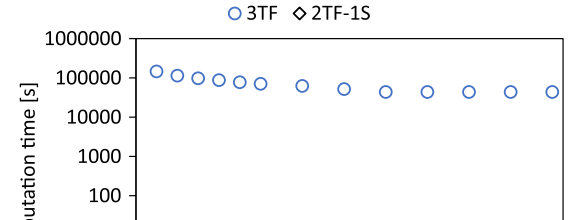

Fig. 11. Dependency of computation time on a with $N = 40$ and $\epsilon = 0$.

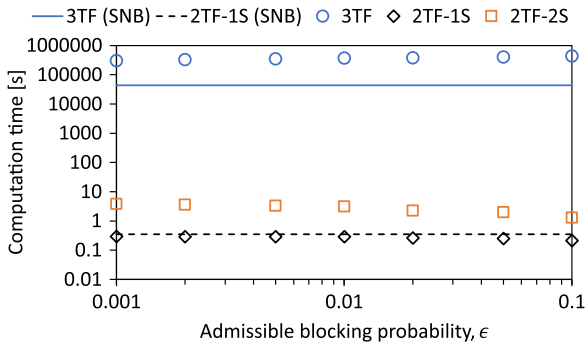

Fig. 12. Dependency of computation time on ϵ with $N = 40$, $a = 2400$, and $p = 0.5$.

We evaluate the computation time along with varying p with $N = 40$, $a = 2400$, and $\epsilon = 0.1$, as shown in Fig. 13. The computation time of 3TF tends to decrease as the value of p is close to 0 or 1. When p takes a value close to 0 or 1, the value of the left side of Eq. (3h) increases, and fewer sets of variables satisfy the equation. Therefore, the number of *for* loops is reduced, and the computation time is shortened. In 2TF-1S and 2TF-2S, the larger p is, the greater the number of *for* loops to find variables that satisfy Eq. (2d) is, so the computation time increases.

C. Number of Used Switches and Number of Used Fibers Depending on the Network Radix

The cost of deploying a switching network increases as the number of used switches and the number of used fibers increase. We investigate the number of used switches and the

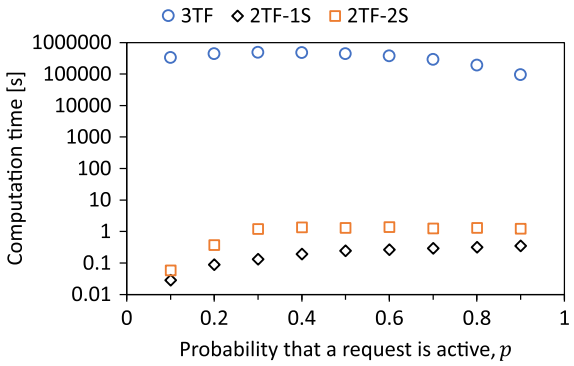

Fig. 13. Dependency of computation time on p with $N = 40$, $a = 2400$, and $\epsilon = 0.1$.

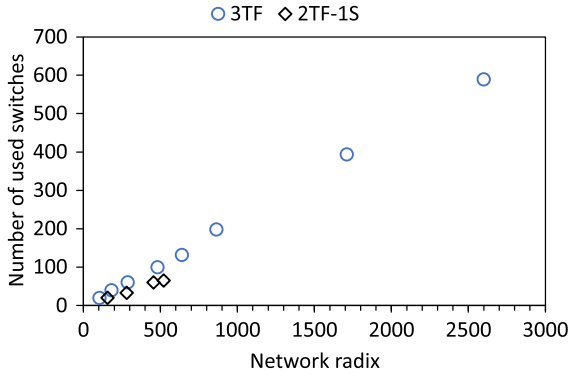

Fig. 14. Number of switches used depending on the network radix with $N = 40$ and $\epsilon = 0$.

number of used fibers depending on the network radix. The number of switches used by 2TF-1S is given by $k + m$, and that of 3TF is given by $k(k' + m') + m$. The number of used fibers of 2TF-1S is given by $2(n + vm)k$, and that of 3TF is given by $(2(n + v'm')k' + 2zm')k + 2vkm$. We compute these numbers from the obtained set of decision variables.

Figure 14 shows the number of switches versus network radix with $N = 40$ and $\epsilon = 0$. We plot the results solved by each design problem using different values of a. When the network radix is up to the maximum value of the network radix of 2TF-1S, which is 520, the number of switches of 2TF-1S is smaller than that of 3TF. On the other hand, 2TF-1S cannot construct a structure when the network radix is larger than the maximum value of the network radix of 2TF-1S.

Figure 15 shows the number of fibers versus network radix with $N = 40$ and $\epsilon = 0$. We plot the results solved by each design problem using different values of a. It shows the same trend as Fig. 14. When the network radix is up to the maximum value of the network radix of 2TF-1S, which is 520, the number of fibers of 2TF-1S is smaller than that of 3TF.

These observations in Figs. 14 and 15 suggest that we choose 2TF if the required network radix does not exceed the maximum value of the network radix of 2TF, and 3TF otherwise.

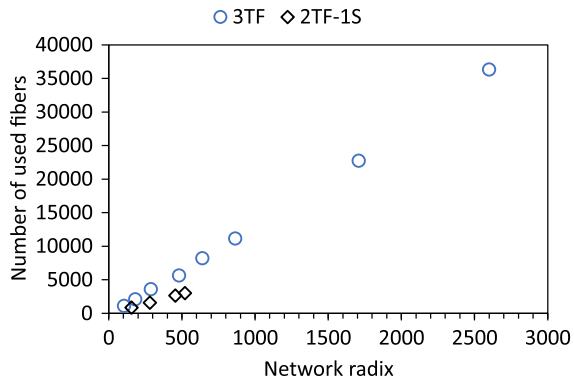

Fig. 15. Number of fibers used depending on the network radix with $N = 40$ and $\epsilon = 0$.

6. APPLICABILITY OF THE PROPOSED MODEL

In the proposed 3TF design model, the 3TF architecture facilitates the passage of an optical circuit connection through five optical switches, connecting a transmitter to a receiver in a data center network. As an illustrative example of its deployability, the 3TF configuration is adaptable for ToRs or aggregation switches as terminals.

We describe examples of power management in an OCS 3TF network configured by the proposed design model. This configuration involves utilizing a pair of transmitters and receivers with an 8 dB power budget [36,37], along with a software-defined optical automatic patch panel [38] serving as each optical switch, ensuring a maximum insertion loss of 1 dB. Additionally, the model suggests employing a piezoelectric-driven optical switch with a switch size of $N \leq 32$ [39], featuring an insertion loss of at most 1 dB for enhanced flexibility.

7. CONCLUSION

This paper has proposed a 3TF design model for OCS with a blocking probability guarantee. The proposed 3TF design model also supports the case of the SNB condition. The proposed model increases the number of stages by making a set of some first-stage switches and some second-stage switches a giant switch in the input-output layer to increase the network radix, i.e., the number of terminals, each with a transmitter and receiver pair connected to the network. We formulated the model as an optimization problem and solved the problem by using an algorithm based on an exhaustive search. Numerical results showed that the proposed model for 3TF has a larger network radix than 2TF in both cases that the proposed model imposes the SNB condition and that it does the admissible blocking probability guarantee. The proposed model for 3TF has a larger increase in network radix when $\epsilon = 0$ to $\epsilon > 0$ than 2TF.

The remaining challenges related to the proposed model include reducing the computation time and handling the fluctuation of p, which is the probability that a request is active at each input port. First, the computation time of 3TF increases

more sharply than that of 2TF in N. Developing a faster algorithm for designing the 3TF, even for larger values of N, poses a challenging task. Second, in our proposed model, the value of p is a hard limit that a request arrival probability does not exceed to guarantee the admissible blocking probability. The request arrival probability can fluctuate and be uncertain. Adopting robust optimization techniques [40–42] can be a candidate for handling such uncertainty in the 3TF design model. This is left for a further study.

Acknowledgment. We are grateful to Haruto Taka for helpful discussions.

REFERENCES

1. H. Liu, R. Urata, K. Yasumura, *et al.*, "Lightwave fabrics: at-scale optical circuit switching for datacenter and machine learning systems," in *Proceedings of the ACM SIGCOMM Conference* (Association for Computing Machinery, 2023), p. 499–515.
2. K.-I. Sato, "Optical switching will innovate intra data center networks [Invited Tutorial]," J. Opt. Commun. Netw. **16**, A1–A23 (2024).
3. L. Poutievski, O. Mashayekhi, J. Ong, *et al.*, "Jupiter evolving: transforming Google's datacenter network via optical circuit switches and software-defined networking," in *SIGCOMM* (2022), pp. 66–85.
4. Y. Mori and K.-I. Sato, "High-port-count optical circuit switches for intra-datacenter networks," J. Opt. Commun. Netw. **13**, D43–D52 (2021).
5. C. Clos, "A study of non-blocking switching networks," Bell Syst. Tech. J. **32**, 406–424 (1953).
6. T. Bowers, "Blocking in 3-stage 'folded' switching arrays," IEEE Trans. Commun. Technol. **13**, 14–37 (1965).
7. A. Jajszcyk, "On combinatorial properties of broadband time-division switching networks," Comput. Netw. ISDN Syst. **20**, 377–382 (1990).
8. W. Kabaciński, *Nonblocking Electronic and Photonic Switching Fabrics* (Springer, 2005).
9. T. Mano, T. Inoue, K. Mizutani, *et al.*, "Increasing capacity of the Clos structure for optical switching networks," in *IEEE Global Communications Conference* (2019).
10. T. Mano, T. Inoue, K. Mizutani, *et al.*, "Redesigning the nonblocking Clos network to increase its capacity," IEEE Trans. Netw. Serv. Manag. **20**, 2558–2574 (2023).
11. E. Oki, H. Taka, and T. Inoue, "Enhancing capacity of optical circuit switching Clos network in data center: progress and challenges," in *33rd International Conference on Computer Communications and Networks (ICCCN)* (2024).
12. Y. Yang and N. H. Kessler, "Modelling the blocking behaviour of Clos networks," Int. J. Parall. Distrib. Syst. Netw. **2**, 1–9 (1999).
13. H. Taka, T. Inoue, and E. Oki, "Design of twisted and folded Clos network with guaranteeing admissible blocking probability," IEEE Netw. Lett. **5**, 265–269 (2023).
14. H. Taka, T. Inoue, and E. Oki, "Twisted and folded Clos-network design model with two-step blocking probability guarantee," IEEE Netw. Lett. **6**, 60–64 (2024).
15. H. Taka, T. Inoue, and E. Oki, "Design model of a twisted and folded Clos network with multi-step grouped intermediate switches guaranteeing admissible blocking probability," J. Opt. Commun. Netw. **16**, 328–341 (2024).
16. E. Oki, Z. Jing, R. Rojas-Cessa, *et al.*, "Concurrent round-robin-based dispatching schemes for Clos-network switches," IEEE/ACM Trans. Netw. **10**, 830–844 (2002).
17. Z. Zheng, L. Qiao, Q. Chen, *et al.*, "A concurrent round-robin-based variable-length packets dispatching scheme for satellite onboard Clos-network switches," in *8th International Congress on Image and Signal Processing* (IEEE, 2015), pp. 1510–1514.
18. W. Sehery and T. C. Clancy, "Load balancing in data center networks with folded-Clos architectures," in *Proceedings of the 1st IEEE Conference on Network Softwarization* (IEEE, 2015).
19. S. Ohta and D. Miyamoto, "Implementation and experimental evaluation of the rebalancing algorithm for folded Clos networks," in *32nd International Telecommunication Networks and Applications Conference* (2022), pp. 312–315.
20. D. Li, M. Xu, Y. Liu, *et al.*, "Reliable multicast in data center networks," IEEE Trans. Comput. **63**, 2011–2024 (2013).
21. C. Camarero, C. Martínez, and R. Beivide, "On random wiring in practicable folded Clos networks for modern datacenters," IEEE Trans. Parallel Distrib. Syst. **29**, 1780–1793 (2018).
22. C. Kachris and I. Tomkos, "A survey on optical interconnects for data centers," IEEE Commun. Surv. Tutorials **14**, 1021–1036 (2012).
23. S. Basu, C. McArdle, and L. P. Barry, "Scalable OCS-based intra/inter data center network with optical ToR switches," in *18th International Conference on Transparent Optical Networks* (2016).
24. Y. Tang, H. Guo, Y. Zhu, *et al.*, "Effectively reconfigure the optical circuit switching layer topology in data center network by OCBridge," J. Lightwave Technol. **37**, 897–908 (2018).
25. P. Li, X. Yu, H. Gu, *et al.*, "FlexNet: an optical switching architecture for optical data center networks," in *19th International Conference on Optical Communications and Networks* (IEEE, 2021).
26. F. P. Kelly, "Blocking probabilities in large circuit-switched networks," Adv. Appl. Probab. **18**, 473–505 (1986).
27. X. Zhao, A. Vahdat, and H. Liu, "Implementation of a large-scale multi-stage non-blocking optical circuit switch," U.S. patent 9,210,487 (12 August 2015).
28. M. Ghobadi, "Emerging optical interconnects for AI systems," in *Optical Fiber Communication Conference (OFC)* (2022), paper Th1G.1.
29. W. Kabaciński, J. Kleban, M. Michalski, *et al.*, "Strict-sense nonblocking networks with three multiplexing and switching levels," in *International Symposium on Networks, Computers and Communications* (2015).
30. M. Głąbowski and M. Sobieraj, "Point-to-group blocking probability in switching networks with multi-service sources and bandwidth reservation," in *9th International Symposium on Communication Systems, Networks and Digital Sign* (2014), pp. 93–98.
31. V. Abramov, S. Li, M. Wang, *et al.*, "Computation of blocking probability for large circuit switched networks," IEEE Commun. Lett. **16**, 1892–1895 (2012).
32. E. Oki, N. Kitsuwan, and R. Rojas-Cessa, "Performance analysis of Clos-network packet switch with virtual output queues," IEICE Trans. Commun. **E94-B**, 3437–3446 (2011).
33. E. Oki, N. Kitsuwan, and R. Rojas-Cessa, "Analysis of space-space-space Clos-network packet switch," in *Proceeding 18th International Conference on Computing and Communication Networks* (IEEE, 2009).
34. D. R. B. de Araújo, C. J. A. Bastos-Filho, and J. F. Martins-Filho, "Methodology to obtain a fast and accurate estimator for blocking probability of optical networks," J. Opt. Commun. Netw. **7**, 380–391 (2015).
35. J. Lin, T. Chang, Z. Zhai, *et al.*, "Wavelength selective switch-based Clos network: blocking theory and performance analyses," J. Lightwave Technol. **40**, 5842–5853 (2022).
36. "JNP-QSFP-100G-PSM4," https://edgeoptic.com/products/juniper/compatible-juniper-100g-optics/jnp-qsfp-100g-psm4/.
37. "QSFP-100G-PSM4," https://edgeoptic.com/products/arista/qsfp-100g-psm4-2/.
38. "G4 NTM," https://www.telescent.com/products.
39. "Polatis 6000i data sheet," https://www.sevensix.co.jp/wp-content/uploads/Polatis_6000i_Data_Sheet.pdf.
40. M. Johnston, H.-W. Lee, and E. Modiano, "A robust optimization approach to backup network design with random failures," IEEE/ACM Trans. Netw. **23**, 1216–1228 (2014).
41. F. He, T. Sato, B. C. Chatterjee, *et al.*, "Robust optimization model for primary and backup resource allocation in cloud providers," IEEE Trans. Cloud Comput. **10**, 2920–2935 (2022).
42. R. Kang, F. He, and E. Oki, "Robust virtual network function allocation in service function chains with uncertain availability schedule," IEEE Trans. Netw. Serv. Manage. **18**, 2987–3005 (2021).

Ryotaro Taniguchi is pursuing the M.E. degree in the Graduate School of Informatics, Kyoto University, Kyoto, Japan. He received the B.E. degree from the Undergraduate School of Electrical and Electronic Engineering, Kyoto University, Japan, in 2024. His research interests include network design, optimization, and algorithm design in data centers.

Takeru Inoue is a Distinguished Researcher at Nippon Telegraph and Telephone Corporation (NTT) Laboratories, Japan. He received the B.E. and M.E. degrees in engineering science and the Ph.D. degree in information science from Kyoto University, Japan, in 1998, 2000, and 2006, respectively. In 2000, he joined NTT Laboratories. From 2011 to 2013, he was an ERATO Researcher with the Japan Science and Technology Agency, where his research focused on algorithms and data structures. His research interests widely cover algorithmic approaches in communication networks.

Kazuya Anazawa is a researcher at Nippon Telephone and Telegraph (NTT) Corporation Network Innovation Laboratories, Japan. He received his B.E. (2016) and M. E. (2018) degrees in computer science and engineering from the University of Aizu. In 2018, he joined NTT Network Innovation Laboratories. His research interests include optical network design and autonomous control of optical networks. He serves as a technical co-lead of the Open Optical & Packet Transport Network Operating System (OOPT-NOS) group in the Telecom Infra Project.

Eiji Oki is a Professor at Kyoto University, Kyoto, Japan. He was with Nippon Telegraph and Telephone Corporation (NTT) Laboratories, Tokyo, from 1993 to 2008, and The University of Electro-Communications, Tokyo, from 2008 to 2017. From 2000 to 2001, he was a Visiting Scholar at Polytechnic University, Brooklyn, New York. His research interests include routing, switching, protocols, optimization, and traffic engineering in communication and information networks.