IEEE Access
Multidisciplinary ⁝ Rapid Review ⁝ Open Access Journal

Received August 17, 2018, accepted September 25, 2018, date of publication October 4, 2018, date of current version October 25, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2873621

Adaptability Analysis for IP Switching and Optical Switching in Geographically Distributed Inter-Datacenter Networks

YANXIA TAN[1], (Student Member, IEEE), RENTAO GU[1], (Member, IEEE), YUEFENG JI[1], (Senior Member, IEEE), DAJIANG WANG[2], AND HONGBIAO LI[2]

[1]Beijing Laboratory of Advanced Information Network, Beijing University of Posts and Telecommunications, Beijing 100876, China
[2]ZTE Corporation, China

Corresponding author: Yuefeng Ji (jyf@bupt.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61871051, in part by the National Science and Technology Major Project under Grant 2017ZX03001016, in part by the Beijing Natural Science Foundation under Grant 4182040, and in part by the ZTE Research Fund.

ABSTRACT The unexpected growth of cloud services lead to a dramatic increase in the network capital expenditure (CAPEX) and latency in geographically distributed inter-datacenter networks. In this paper, an adaptability analysis is proposed on CAPEX and forwarding delay for IP switching and optical transport network (OTN) switching in distributed inter-datacenter networks. By abstracting the switching behavior of the IP routers and OTN equipment, the IP nodes, and OTN nodes are mapped to the switching matrix of CLOS switching network. A unified network-level connection model is established based on CLOS switching network to meet the needs of connectivity. Then, the hardware costs and the forwarding delay are evaluated to find out the bandwidth threshold (BDT) which decides the most suitable switching mode for a specific switching granularity. The adaptability of IP switching and OTN switching under different switching granularities and port rates is performed based on the simulation platform in terms of hardware cost, optical port cost, and forwarding delay. Simulation results show that the BDT based on hardware cost is 280 Mbps when switching granularity is 1.25 Gb/s. For 2.5 Gb/s switching granularity, when average bandwidth between multiple datacenters is higher than 600 Mbps, OTN switching has more advantages than IP switching on capital expenditure, especially when port rate of the equipment is high. In addition, the simulation results also verify that the forwarding delay is reduced obviously with OTN switching.

INDEX TERMS CAPEX, costs, distributed inter-datacenter network, forwarding delay, OTN switching, optical interconnections, packet switching.

I. INTRODUCTION

Toward 5G and beyond, driven by the growing cloud service applications such as 4K video, VR/AR and content delivery services, the global Internet traffic has increased rapidly [1]. Large enterprises, such as Google and Facebook, are adopting geographically distributed (geo-distributed) datacenters (DCs) to ensure high-quality and reliable services for the world-wide customers [2]. In a geo-distributed DC system, the exponential traffic growth may lead to a dramatic increase in the network capital expenditure (CAPEX) cost. Furthermore, the emerging technologies such as big-data analytics, live-TV, etc. have higher requirements on latency [3], [4]. The delay-sensitive datacenter services must be exchanged in real-time under rigid delay bounds. Therefore, current datacenter interconnection networks need to evolve towards a flexible, low-cost and low-latency network platform.

Current datacenter interconnection networks are divided into independent IP/MPLS networks (IP layer) and optical transport networks (optical layer) [5], [6] to meet the traffic growth. Generally, the electrical IP layer switching is the core technology to meet the bandwidth requests, and optical layer only provides a quasi-static end-to-end physical link for the IP layer in service provisioning. Hence the data transmission in a geo-distributed DC network requires multiple electrical-to-optical (E/O) and optical-to-electrical (O/E) conversions. A study shows that 55%-85% of the service requests processed by an IP node are just passing through to another IP node [7], and the service requests do not originate or terminate at this location. Transit traffic consumes 50%-80% of resources in IP routers [8]. Moreover, the links in IP networks today are usually overprovisioned to deal with traffic fluctuation [5], which greatly increases the network CAPEX cost.

To decrease the CAPEX and latency, the multi-layer datacenter interconnection network is well studied, and several methods have been proposed in the past decades. Usually, the existing studies on the cost problem for inter-datacenter networks are confronted as an integer linear programming (ILP) or mixed integer linear programming (MILP) optimization problem [5], [9], which are subject to different types of restrictions [10]. Heuristic approaches are proposed to save cost which include 1) increase the infrastructure capacity [11], 2) keeping the traffic in the optical domain (via optical bypass) instead of letting it touch IP router [12]–[14], and finally 3) performing packet and optical layer joint optimization with the help of software-defined networking (SDN) [15] technology to reduce cost [16]–[19]. However, increasing the capacity of their infrastructure to cope with the exponential traffic growth will greatly increase the cost of network construction [11], [19]. And reference [19] shows that optical bypass results in only 10%-15% CAPEX cost saving. Furthermore, in the process of joint optimization, adjusting a link will cause impact on other links, nodes, and even the entire network, which requires multiple optimizations to achieve the goal [6], [17], [19].

Currently, service requests with small bandwidth and low latency requirements are processed in the edge micro-DCs (mDC) [20], [21], and the service requests interacting among multi-datacenters (multi-DCs) often shows high-bandwidth characteristics. Therefore, fine-grained switching at the inter-datacenter network has diminished gradually, and optical layer switching with the advantages of large capacity, high bandwidth, and low latency may have more advantages on cost and latency over IP layer switching. A promising approach to achieve significant cost saving and latency decreasing in geographically distributed inter-datacenter networks is harnessing optical transport technologies [3], [22] such as Optical Transport Network (OTN) [23], [24] to satisfy the high-speed requests. However, the large switching granularity of optical switching is more likely to reduce the resource utilization and further increase the network capital expenditure, when the service request bandwidth is smaller than the switching granularity. Therefore, the adaptability of IP switching and optical switching under different bandwidth conditions need to be studied.

In this paper, we focus on the adaptability analysis for IP switching and optical switching in geographically distributed inter-datacenter networks by calculating the CAPEX and forwarding delay of the network components. We wish to obtain the bandwidth adaptation range for the two switching modes. The geo-distributed inter-datacenter network, in which multiple geographically distributed DCs are interconnected by OTN equipment is introduced. Next, by abstracting the switching behavior of IP routers and OTN equipment, a unified network-level connection model to calculate the CAPEX and the forwarding delay is established. Then we introduce the adaptability assessment process to determine the bandwidth threshold (BDT) with connectivity guarantee based on CLOS [25] switching network. Finally, under different traffic and network parameters, we analyze the hardware cost, optical port cost and forwarding delay to obtain the adaptation range for IP switching and OTN switching.

The rest of the paper is organized as follows. Section II introduces the geographically distributed inter-datacenter network architecture. The mathematical model for calculating the CAPEX and the forwarding delay of IP switching and OTN switching, and the proposed adaptability assessment process are discussed in Section III. Then, Section IV presents the performance evaluation with simulation results. Finally, we summarize the paper in Section V.

II. DISTRIBUTED INTER-DATACENTER NETWORK ARCHITECTURE AND COSTS

The multi-layer inter-datacenter network architecture usually consists of an optical transport network and several packet networks. The infrastructure of the transport layer can be optical transport network, wavelength division multiplex (WDM), elastic optical network (EON) or a combination of these networks [18]. In this paper, we assume that the network infrastructure is based on OTN [23] technology.

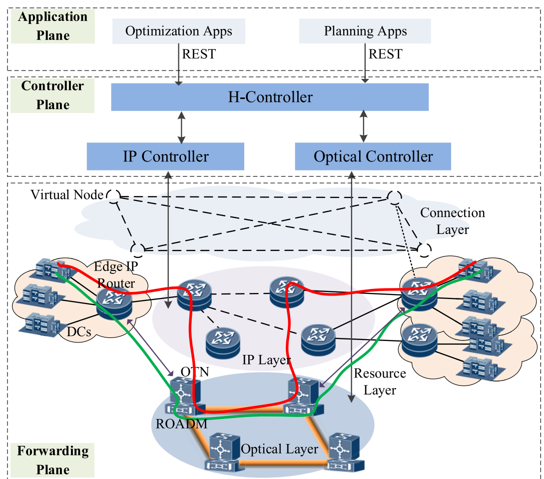


FIGURE 1. Architecture of the geographically distributed inter-datacenter network.

A. DISTRIBUTED INTER-DATACENTER NETWORK ARCHITECTURE

A geographically distributed datacenter architecture is shown in Fig. 1, which is composed of three planes: 1) application plane, 2) controller plane, and 3) forwarding plane. The application plane includes various optimization applications (APPs) and planning APPs, and it operates the underlying forwarding plane by calling the control plane. The control plane provides resource management for the IP routers and the OTN equipment through the hierarchical

SDN controller (H-Controller) [15]. For the forwarding plane, multiple geographically distributed DCs are interconnected by multi-domain networks to accommodate the service requests according to the strategies provided by the control plane. The connection layer concerns the reachability of traffic flows while the resource layer dynamically adjusts the resources to satisfy the service requests.

In the process of service transmission, edge IP router nodes are mapped onto the virtual nodes of connection layer when converging the service requests from different DCs. In networks, the data transmission requires multiple E/O and O/E conversions depicted as the red line shown in Figure 1, which greatly increase the network CAPEX and forwarding delay. Therefore, reducing the processing times of the IP routers will be a promising approach to reduce the cost and forwarding delay in geo-distributed inter-datacenter networks. Optical layer switching with the advantages of large capacity, high bandwidth and low latency may have more advantages on cost and latency over IP layer switching. However, the large switching granularity of optical switching is more likely to reduce the resource utilization and further increase the network capital expenditure, when the service request bandwidth is smaller than the switching granularity. Hence, when the bandwidth exceeds a certain threshold, the service requests can be directly transmitted through the OTN equipment as the green line shown in Figure 1, which saves the processing time of the IP routers and thereby reduce the forwarding delay. On the contrary, when the bandwidth is lower than the threshold, the service requests should be transmitted through the IP router and OTN equipment. How to evaluate the bandwidth threshold is an important issue need to be investigated.

To assess the bandwidth adaptability of IP switching and optical switching, we adopt OTN switching as an example. If the service requests are directly transmitted through the OTN equipment, the OTN equipment in resource layer is divided into the convergence layer equipment and the core layer equipment [26]. The OTN equipment in convergence layer performs optical and electrical hybrid processing through the electrical cross matrix, and the OTN equipment in core layer performs full optical processing through Reconfigurable Optical Add-Drop Multiplexer (ROADM) to ensure high-quality and reliable services for the customers. Correspondingly, the IP network is also defined as a hierarchical architecture. The outer-core router nodes are equipped with converge routers which support aggregating services from distributed DCs, and core nodes are equipped with core routers to provide a second level of traffic converge between source-destination pairs of converge routers. We wish to obtain the bandwidth adaptation range on CAPEX and then evaluate the forwarding delay for the two switching modes.

B. COMPOSITION OF THE NETWORK EQUIPMENT COST

For the convenience of comparison, we take normalized data for simulation verification in this paper. The detailed cost parameters for OTN equipment and IP router are shown in Table 1 and Table 2, and the costs are expressed in normalized values. The composition of the OTN equipment cost includes the common unit and the service unit. Correspondingly, the IP router includes the basic system, the line card and the optical module. The common unit of the OTN equipment and basic system of the IP router are the basic platform for the equipment which are service-independent configuration items. On the contrary, the service unit of the OTN equipment, the line card and optical module of the IP router are service-aware configuration items. One mother card occupies one slot and two sub cards can be inserted in one mother card. For example, for an IP router with two 100G ports, we need one base system, one A-type network processor card, two 100G cards and two 100G optical modules whose total hardware cost is $10+21.14+2\times58.57+2\times53.97 = 256.22$. The detailed latency and latency jitters parameters for OTN equipment and IP router are shown in Table 3.

TABLE 1. Cost parameters for OTN equipment (NORMALIZED).

OTN Equipment	Composition	Cost
Common unit	Common unit	3.83
Service unit	16 port 2.5G client side card	1.25
	12 port 10G client side card	2.22
	2 port 40G client side card	2.29
	2 port 100G client side card	1.59
	1 port 100G optical module	1.83

TABLE 2. Cost parameters for IP router (NORMALIZED).

IP Router	Composition	Cost
Base System	Base System	10
Line card	A-type network processor card	21.14
	1 port 100G card	58.57
	3 port 40G card	52.06
	10 port 10G card	24.57
	B-type network processor card	84.13
	8 port 2.5G card	97.62
Optical module	100G optical module	53.97
	40G optical module	8.41
	10G optical module	0.22
	2.5G optical module	0.16

TABLE 3. Typical latency parameters for OTN equipment and IP router.

Equipment	Latency (μs)	Latency jitters (μs)
OTN equipment	5-10	1
IP Router	10-50	1-10

III. PROBLEM FORMALIZATION

In this section, we introduce the proposed mathematical model to calculate the hardware cost and the forwarding delay of IP switching and OTN switching in distributed inter-datacenter networks and then illustrate the adaptability assessment process.

A. MATHEMATICAL MODEL

Due to the large switching granularity of optical switching, how to ensure the connectivity is a key issue in networking. To ensure the connectivity, the strictly non-blocking CLOS switching network [25] is employed for networking in this paper. CLOS switching network is a kind of widely used multi-level non-blocking network. Under the strict non-blocking conditions, a connection can be established in the switching network at any time if the source node and the destination node of the connection are free, without affecting the established connections in network. In this study, we abstract the switching behavior of IP routers and OTN equipment, and then establish a unified network-level connection model based on CLOS switching network to ensure the connectivity of the networks.

In the unified network-level model, a strict non-blocking three-level symmetric CLOS switching network is constructed, and then split it into a multi-level CLOS switching network based on the service and network parameters. Next, we determine how many routers or OTN equipment we needed to transmit the datacenter services, and then calculate the number of links which make up the edge of networks. In the models, one switching matrix (SM) of the multi-level CLOS switching network is mapped to an OTN equipment or an IP router, and the connection between SMs represents the fiber link in geo-distributed datacenter networks correspondingly as shown in Figure 2. Service requests across multiple DCs, from a server in one DC to another in a different DC are interconnected by the OTN transport networks.

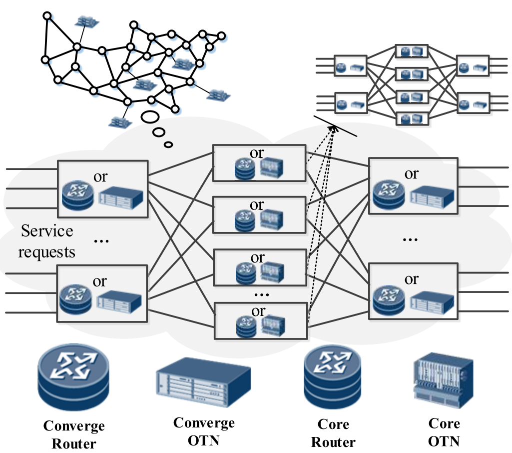


FIGURE 2. Node and link mapping based on CLOS switching network.

The network topology is represented by a connected graph $G(V_1, V_2, E)$, where V_1 represents the set of convergence layer physical nodes, V_2 represents the set of core layer physical nodes, and E represents the set of physical links between node pairs. OTN defines the Optical channel Data Unit-k (ODUk) time division multiplexing sublayer, which supports multiplexing several lower bit-rate signals into a higher bit-rate signal. In the rest of the paper, U_k denotes ODUk (k = 0, 1, 2, 3, 4). Then for the multiplexing process, we can write: $2U_0{=}U_1, 4U_1{=}U_2, 4U_2{=}U_3, 2U_2{=}U_4$. To describe the mathematical model, a set of required notations, costs and variables are defined and described in Tables 4. We assume that one core equipment has eight mother card, i.e. $N^j_{ip\_c} = 8$ and $N^j_{o\_c} = 8$.

Assuming that the bandwidth of the service request Λ_{sd} across multiple DCs follows the normal distribution $N(\mu_{sd}, \sigma^2_{sd})$, the average bandwidth value is μ_{sd}, the standard deviation value is σ^2_{sd}, and the probability density function is $f(x)$. The peak bandwidth b^α_{sd} of the service request is obtained from the value of the allowable blocking rate $\alpha{=}\int_{b^\alpha_{sd}}^{+\infty} f(x)dx$. In this work, OTN equipment are overprovisioned to account the uncertain traffic flows such as the sudden traffic surges in networks. Peak rate requests will be successfully attended by the available OTN equipment in the module. The objective function captures the total hardware cost of the network over convergence layer and core layer based on the premise of ensured connectivity. The total hardware cost is defined by Eq. (1) and Eq. (2).

TABLE 4. Notations, costs and variables.

Notation	Description
N_r	Total number of service requests
$P = \{1, 2, 3, 4\}$	The card type set for IP router / OTN equipment
C_j	Port rate set for different types of card
$N^j_{ip\_c}/N^j_{o\_c}$	Number of mother card j on one IP router / OTN equipment
$N^j_{ip\_p}/N^j_{o\_p}$	Number of port for IP router / OTN equipment on one card j
Cost Parameter	**Description**
C_{ip_b}	Cost for base system of IP router
$C^j_{ip\_A}/C^j_{ip\_B}$	Cost for A-type / B-type card of IP router
$C^j_{ip\_i}/C^j_{ip\_m}$	Cost for line card / optical module of IP router
C_{o_c}	Cost for common unit of OTN equipment
$C^j_{o\_b}/C_{o\_m}$	Cost for service unit / optical module of OTN equipment
Variable	**Description**
C_{t_ip}/C_{t_otn}	Total hardware cost for IP router / OTN equipment
N^r_e/N^r_c	Number of convergence / core IP router for Λ_{sd}
N^o_e/N^o_c	Number of convergence / core OTN equipment for Λ_{sd}
B_{sd}	The bandwidth for Λ_{sd} after the OTN multiplexing
n_{iter_r}/n_{iter_o}	The iteration time for split the middle layer SM for IP switching / OTN switching
$N^j_{p\_ip}/N^j_{p\_o}$	Number of available sub card for one IP router / OTN equipment under strict non-blocking constraint

$$C_{t\_ip} = 2 \cdot N^r_e \cdot C_{ip} + N^r_c \cdot C_{ip} \quad (1)$$

$$C_{t\_otn} = 2 \cdot N^o_e \cdot C_{otn} + N^o_c \cdot C_{otn} \quad (2)$$

Where Eq. (1) and Eq. (2) refers to the total hardware cost of IP switching and OTN switching, respectively. Optical switching has fixed switching granularity, so the number of service requests on per port is fixed. In optical transport networks, service requests should be adapted to a fixed ODUk granularity. OTN layer adaptation constraints are evaluated as:

$$B_{sd} = U_k, U_{k-1} < b^\alpha_{sd} \le U_k, \quad \forall sd \in r, U_{-1} = 0 \quad (3)$$

Strict Non-blocking constraints are shown below. These constraints assure that a service request can be established in networks at any time if the source server and the destination server of the request are free.

$$N^j_{p\_ip} = \left\lceil (2 \cdot N^j_{ip\_c} + 1)/2 \right\rceil, \quad \forall j \in P \quad (4)$$

$$N^j_{p\_o} = \left\lceil (2 \cdot N^j_{o\_c} + 1)/2 \right\rceil, \quad \forall j \in P \quad (5)$$

We assume that the first three types of line card need A-type network processor card and the fourth type needs

B-type network processor card for IP router. Costs for single IP router C_{ip} and single OTN equipment C_{otn} are evaluated as:

$$C_{ip} = \sum_{j=1}^{3} N_{ip\_c}^{j} \cdot (C_{ip\_A} + 2 \cdot C_{ip\_i}^{j} + 2 \cdot N_{ip\_p}^{j} \cdot C_{ip\_m}^{j}) + N_{ip\_c}^{4} \cdot (C_{ip\_B} + 2 \cdot C_{ip\_i}^{4} + 2 \cdot N_{ip\_p}^{4} \cdot C_{ip\_m}^{4}) + C_{ip\_b} \quad (6)$$

$$C_{otn} = 2 \cdot \sum_{j \in P} N_{o\_c}^{j} \cdot (C_{o\_b} + \left\lfloor C_j \cdot N_{o\_p}^{j} / C_4 \right\rfloor \cdot C_{o\_m}) + C_{o\_c} \quad (7)$$

After splitting the three-level CLOS switching network, the number of convergence layer equipment are evaluated as:

$$N_e^r = \left\lfloor N_r / N_{p\_ip}^{j} \cdot (C_j \cdot N_{ip\_p}^{j} / \mu_{sd}) \right\rfloor, \quad \forall j \in P \quad (8)$$

$$N_e^o = \left\lfloor N_r / N_{p\_o}^{j} \cdot (C_j \cdot N_{o\_p}^{j} / B_{sd}) \right\rfloor, \quad \forall j \in P \quad (9)$$

Core layer equipment number is evaluated as follows, where $r_{ip}(t)/r_o(t)$ represents the number of SMs on per first-layer, and $n_{c\_r}(t)/n_{c\_o}(t)$ represents the number of SMs for per middle-layer when splitting the CLOS switching network for IP switching and OTN switching, respectively. The formula for calculating the value of $n_{c\_o}(t)$ is similar with $n_{c\_r}(t)$.

$$r_{ip}(1) = \left\lfloor \frac{N_e^r \cdot (C_j \cdot N_{ip\_p}^{j} / \mu_{sd})}{N_{p\_ip}^{i} \cdot (C_i \cdot N_{ip\_p}^{i} / \mu_{sd})} \right\rfloor, \quad \forall i, j \in P \quad (10)$$

$$r_{ip}(t) = \left\lfloor r(t-1) / N_{p\_ip}^{i} \right\rfloor, \quad t = 2, 3, \ldots, n_{iter\_r}, \ i \in P \quad (11)$$

$$n_{c\_r}(n_{iter\_r}) = 2 \cdot N_{ip\_c}^{i} \cdot n_{c\_r}(n_{iter\_r} - 1) + 2 \cdot r_{ip}(n_{iter\_r}), \quad n_{c\_r}(0) = 1, \ i \in P \quad (12)$$

The number of core layer IP router N_c^r for IP switching and the number of core layer OTN device N_c^o for OTN switching are evaluated as:

$$N_c^r = 2 \cdot N_{ip\_c}^{j} \cdot n_{c\_r}(n_{iter\_r}), \quad \forall j \in P \quad (13)$$

$$N_c^o = 2 \cdot N_{o\_c}^{j} \cdot n_{c\_o}(n_{iter\_o}), \quad \forall j \in P \quad (14)$$

For OTN switching, the total hardware cost is the minimum cost value under the condition of guaranteeing the blocking rate α. According to the total number of service requests and the average bandwidth value, the hardware cost of IP switching and OTN switching are calculated. At the same time, the forwarding delay for IP switching T_r and OTN switching T_o are calculated by Eq. (15) and Eq. (16).

$$T_r = (2 \cdot n_{iter\_r} + 1) \cdot t_r \quad (15)$$

$$T_o = (2 \cdot n_{iter\_o} + 1) \cdot t_o \quad (16)$$

Where t_r and t_o represent the forwarding delay on per node for IP switching and OTN switching, respectively.

Algorithm 1 Network Adaptability Assessment Process

Given: μ_{sd} (Mbps), α, service request number set χ_r;
Output: BDT based on hardware cost ϕ_{bdt_hc};

1: **For** each **t** in χ_r do
2: $\mu_{sd} = 1$;
3: **While** $\mu_{sd} \geq 1$ do
4: Calculate the b_{sd}^{α} according to μ_{sd} and α;
5: Establish a three-level symmetric clos switching network under strict non-blocking principle;
6: Map the first and third layer SM to the convergence layer equipment, and then calculate the N_e^r and N_e^o by Eq.(8) and Eq.(9);
7: Split and map the middle layer SM based on the port rate of core devices;
8: Calculate the C_{ip} and $C_{otn}$by Eq.(6) and Eq.(7);
9: Calculate the N_c^r and N_c^o by Eq.(13) and Eq.(14);
10: Calculate the C_{t_ip} and C_{t_otn} by Eq.(1) and Eq.(2);
11: **If** $C_{t\_ip} \geq C_{t\_otn}$
12: $\phi_{bdt\_hc} = \mu_{sd}$ and $\mu_{sd} = 0$;
13: **else**
14: $\mu_{sd} = \mu_{sd} + 1$.
15: end **If**
16: end **While**
17: end **For**
18: **Return** ϕ_{bdt_hc} for different number of service requests.

B. ADAPTABILITY ASSESSMENT PROCESS

To describe the adaptability assessment process, we introduce the concept of bandwidth threshold (BDT), which is a demarcation point for differentiating the adaptation range for IP switching and OTN switching on average bandwidth under different number of service requests. When the average bandwidth of service requests accesses the bandwidth threshold based on hardware cost (BDT-HC), the cost of using OTN switching is lower than that of using IP switching, and inversely, when the average bandwidth of service requests is lower than the BDT-HC, IP switching shows better performance on hardware cost. The following adaptability assessment process illustrates the steps involved in estimating the bandwidth adaptation range of IP switching and OTN switching.

In the adaptability assessment procedure, the total hardware cost for OTN switching is the minimum cost value under the condition of guaranteeing the blocking rate α, which guarantees there is sufficient capacity to sustain the service request over an extended period of time. For IP switching, we assign port resources for service requests according to the average bandwidth value. By applying the cost parameters shown in Table 1 and Table 2 to the abstracting model, the hardware cost for IP switching and OTN switching under different number of service requests are calculated, and then we found out the BDT-HC for the two switching modes by comparing the cost value.

IV. SIMULATION RESULTS AND PERFORMANCE ANALYSIS

The switching granularity of optical switching is larger than IP switching. Therefore, to ensure the sufficient capacity for

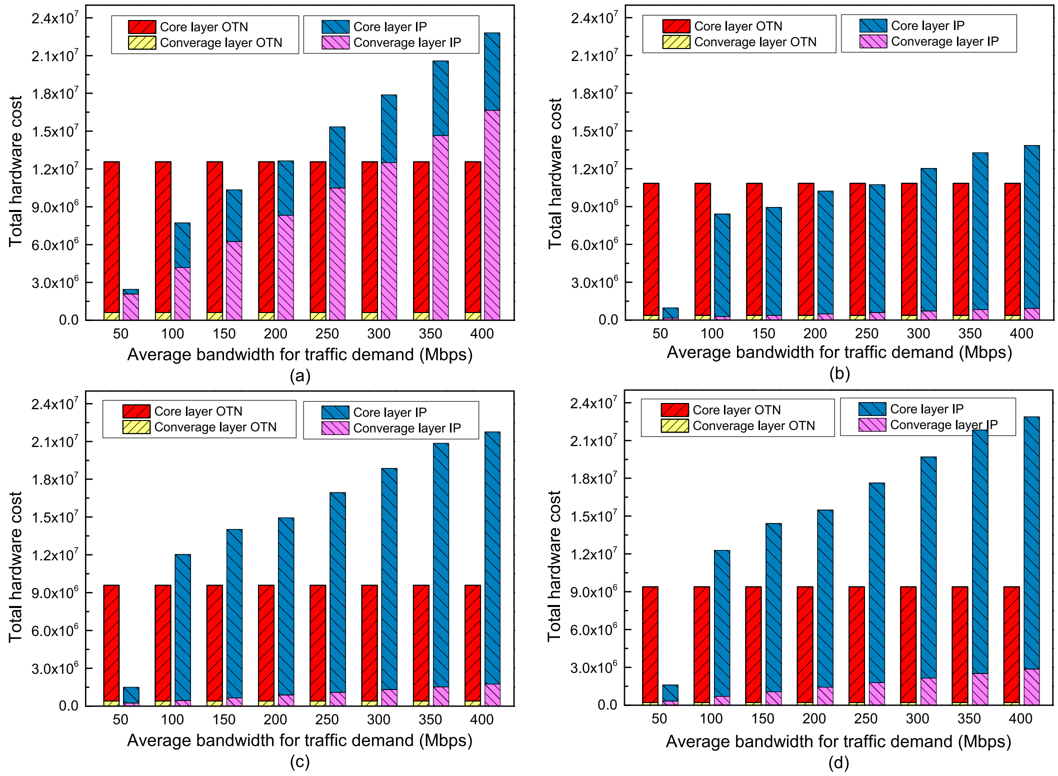


FIGURE 3. Total hardware cost under different port rate (ODU0): (a) 2.5Gbps/10Gbps; (b) 10Gbps/40Gbps; (c) 40Gbps/100Gbps; (d) 100Gbps/100Gbps.

achieving the required number of service requests, the number of optical equipment must increase correspondingly which results in the increase of the optical cables and network hops. Taking OTN technology as an example, we firstly estimate the bandwidth threshold for IP switching and OTN switching with the simulation results of hardware costs, and then the results of the bandwidth threshold are verified on multi-rate scenario. Finally, to verify the number of increased optical cables and network hops are in controllable range for ensuring the network connectivity, the optical ports cost and the forwarding delay are analyzed.

In simulations, IP router and OTN equipment supports four port rates $P_v = [2.5, 10, 40, 100]$ (in Gbps). And the port rate for core layer equipment is greater than the convergence layer equipment. Port rate of 40Gbps/100Gbps indicates that port rate for convergence layer device is 40Gbps, and port rate for core layer devices is 100Gbps. The switching granularity of the transport channels for the simulation experiments are 1.25 Gb/s (ODU0) and 2.5 Gb/s (ODU1). We assume that the bandwidth of the service requests follows the normal distribution, and the standard deviation is 100Mbps. For OTN switching, bandwidth is over-dimensioned for traffic peaks under a certain blocking rate to prevent the burst of random services. Moreover, we assume the allowable blocking rate is 0.01, and the bandwidth of a single wavelength is 100G. The average bandwidth of service requests across multiple DCs is adjusted according to the switching granularity to find out the BDT.

A. BDT BASED ON HARDWARE COST

The BDTs of the IP switching and OTN switching based on hardware cost are investigated in this subsection. Assuming that the number of service requests across geo-distributed DCs is between 1×10^4 and 5×10^6, we compare the BDT-HC when the switching granularity for OTN equipment are changed under different port rates.

We firstly calculate the total hardware cost for the two switching modes when the total number of service requests is 3×10^6. The hardware cost includes convergence layer cost and core layer cost. Figure 3 and Figure 4 show the total hardware cost for OTN switching and IP switching under different port rate when the available switching granularity of OTN equipment are ODU0 and ODU1, respectively. We can see in Figure 3 and Figure 4 that the major hardware cost component for OTN switching is core layer equipment, and the greater the port rate is, the lower the hardware cost is. The lowest total hardware cost is obtained when port rate is 100Gbps/100Gbps, and this indicates that the hardware cost for OTN switching in networks can be reduced when increasing the port rate of OTN equipment. In Figure 3 and Figure 4, the major hardware cost component for IP switching is the convergence layer routers when port rate is 2.5Gbps/10Gbps. And inversely, when port rate becomes high, the core layer routers become the main contribution of the total hardware cost. This is because guaranteeing there is sufficient capacity to sustain the service requests, multiple core IP routers should be deployed in networks. We can see significant cost

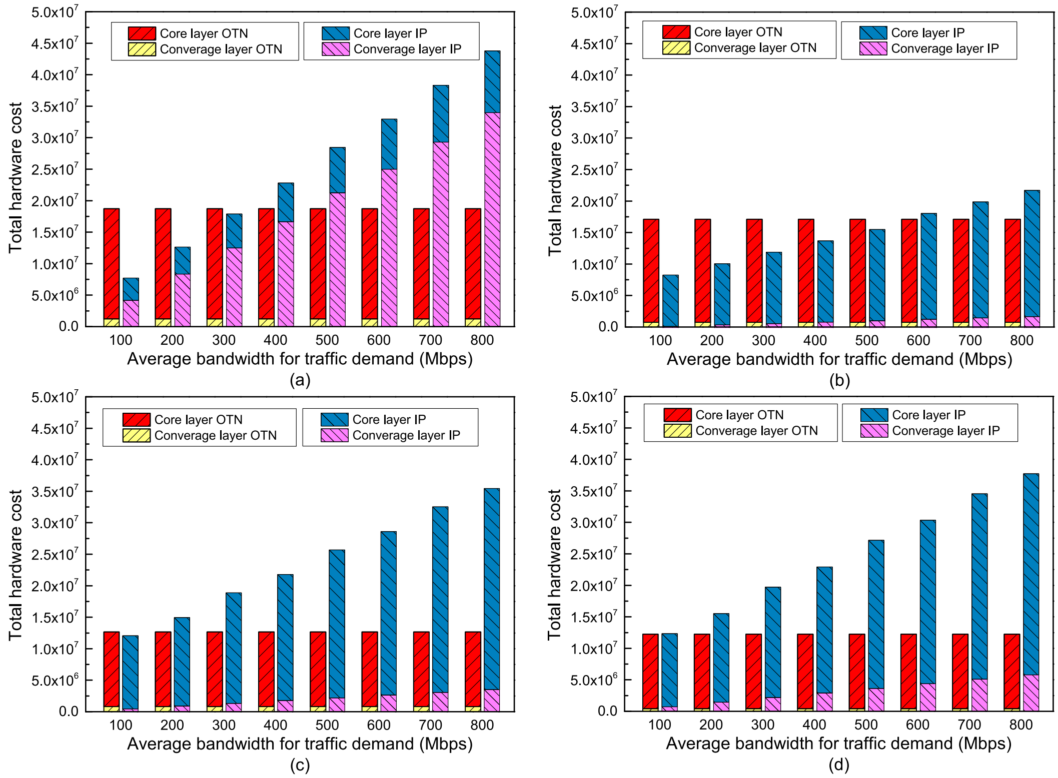


FIGURE 4. Total hardware cost under different port rate (ODU1): (a) 2.5Gbps/10Gbps; (b) 10Gbps/40Gbps; (c) 40Gbps/100Gbps; (d) 100Gbps/100Gbps.

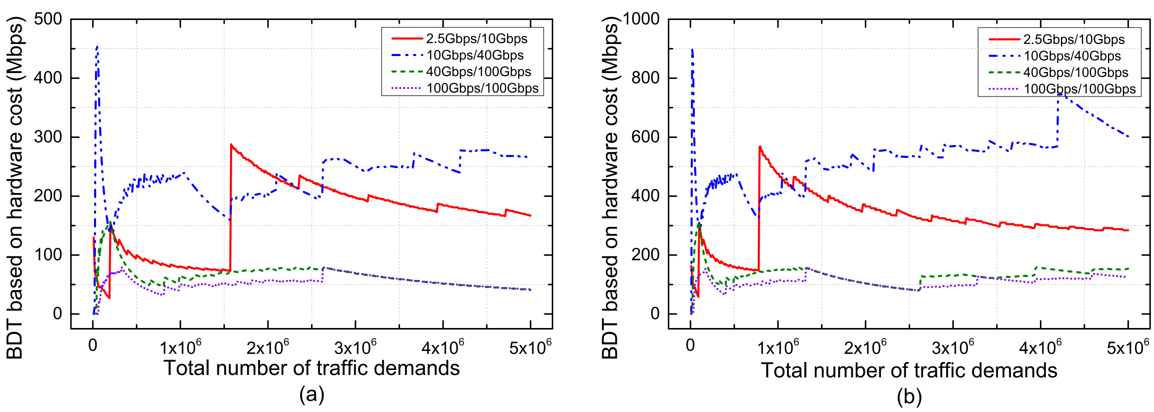


FIGURE 5. BDT-HC under different number of service requests: (a) Switching granularity is 1.25 Gb/s; (b) Switching granularity is 2.5 Gb/s.

reduction for IP switching when port rate for convergence layer IP router is 10Gbps and core layer IP router is 40Gbps in Figure 3(b) and Figure 4(b). It indicates that increasing the port rate for IP routers cannot reduce the hardware cost in networks for IP switching. In addition, we also can observe that the total hardware cost for OTN switching is lower than IP switching when bandwidth for service request excesses a threshold value in Figure 3 and Figure 4.

Next, we further obtain the threshold value for OTN switching and IP switching mentioned above. Figure 5(a) shows the BDT-HC results when switching granularity is 1.25 Gb/s. As can be seen, when port rate is 100Gbps/100Gbps and 40Gbps/100Gbps, the threshold value is about 60Mbps. And the BDT-HC are fluctuated greatly when the number of service requests is small. The reason is that in order to ensure the connectivity, the amount of convergence layer devices varies greatly for OTN switching when the service requests number is small. It can be seen in Figure 5(a) that when the average bandwidth of service requests for geo-distributed datacenters is relatively high, e.g. average bandwidth excessing 280Mbps, the hardware cost of OTN switching is lower in comparison with the hardware cost of the IP switching when switching granularity is 1.25 Gb/s. Furthermore, in Figure 5(b), we report the threshold value when switching granularity is 2.5 Gb/s. The results confirm that when the average bandwidth exceeds 600Mbps,

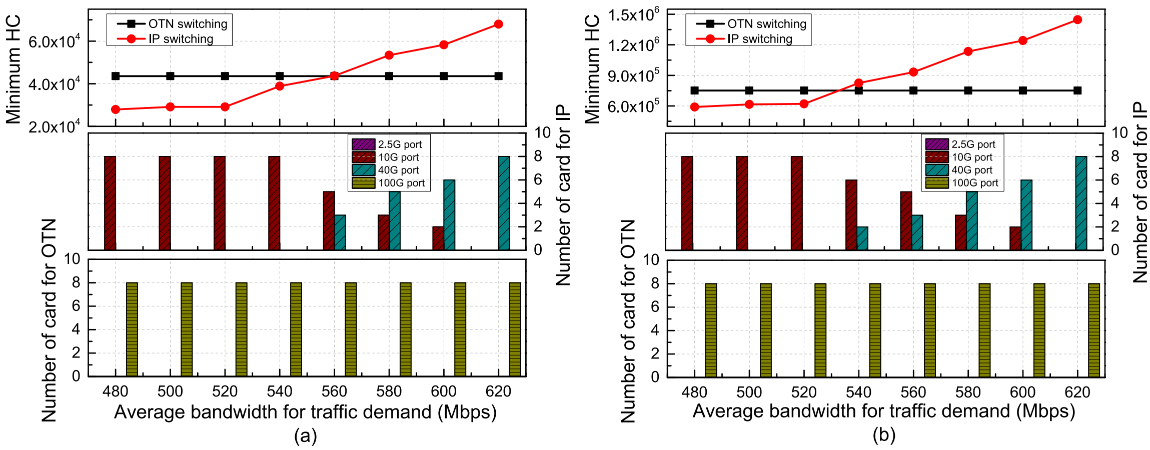


FIGURE 6. **Minimum hardware cost and the corresponding card configuration: (a) Number of service requests is 5×10^4; (b) Number of service requests is 4×10^5.**

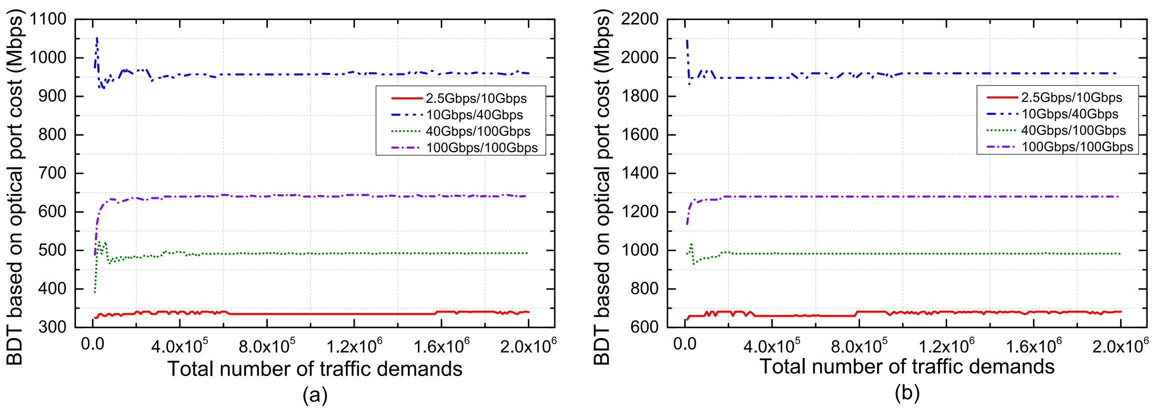


FIGURE 7. **BDT-OPC under different number of service requests: (a) Switching granularity is 1.25 Gb/s; (b) Switching granularity is 2.5 Gb/s.**

OTN equipment can replace the IP routers to serve the service requests, and the larger the port rate is, the more obvious the advantage of OTN switching is on hardware cost under strict non-blocking condition.

From the above analysis, we can know that the great port rate will effectively reduce the hardware cost for the OTN switching. Nevertheless, the cost advantage of the IP switching is not obvious when the port rate becomes high. In this part, we conduct simulation experiments to verify the bandwidth threshold results mentioned above. As the number of service requests increases, the bandwidth threshold value fluctuation shrinks. So we carry out the simulation verification on multi-rate scenario when the number of service requests is 5×10^4 and 4×10^5 (two intermediate values). In multi-rate scenario, the OTN equipment and IP routers support mixed-line-rate (MLR) transmission (2.5/10/40/100 Gbps). Figure 6 shows the minimum hardware cost and the corresponding port card configuration under different average bandwidth. In figure 6(a), hardware cost for IP switching is higher than OTN switching when bandwidth for service request excess 560Mbps. We can see that when all port cards are 100Gbps, OTN switching achieves the minimum hardware cost. And correspondingly, when port cards are 10G or 40G, IP switching achieves the minimum hardware cost, which is consistent with the above analysis results. Likewise, the same conclusion is obtained when the number of service requests is 4×10^5 as shown in figure 6(b), hardware cost for IP switching is higher than OTN switching when average bandwidth for service request exceeding 535Mbps.

B. BDT BASED ON OPTICAL PORT COST

In order to ensure there is sufficient capacity to achieve the required number of service requests, the amount of OTN hardware must be increased. As a result, either amount of optical cables is also increased for OTN switching. We evaluate and analyze the optical cables demand of OTN switching and IP switching by the cost of the needed optical ports. In this study, we mainly observe the bandwidth threshold based on optical port cost (BDT-OPC) under different port rate.

In the simulations, we consider the cases that the optical port for IP switching is twice times more expensive than the OTN switching optical port, because of the resource over-provisioning in IP network. The total number of service requests is from 1×10^4 to 2×10^6. The BDT-OPC performance of IP switching and OTN switching under different number of service requests is shown in Figure 7. We observe

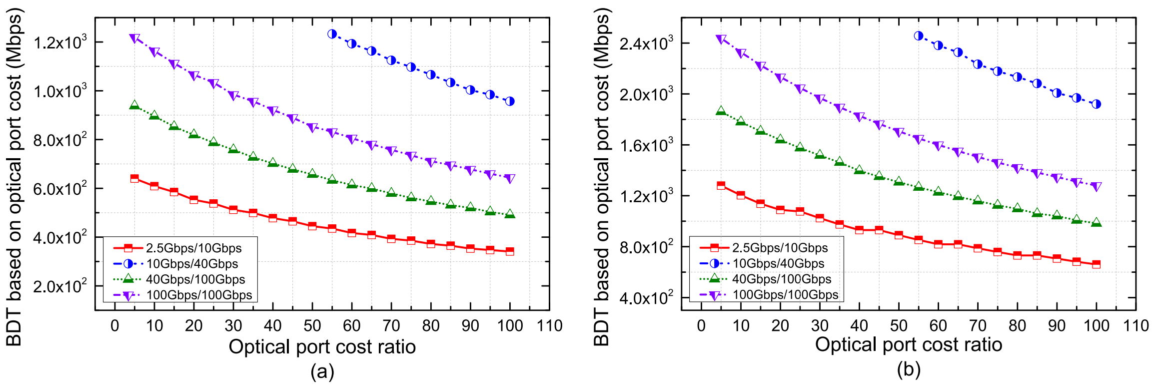


FIGURE 8. BDT-OPC under different optical port cost ratio: (a) Switching granularity is 1.25 Gb/s; (b) Switching granularity is 2.5 Gb/s.

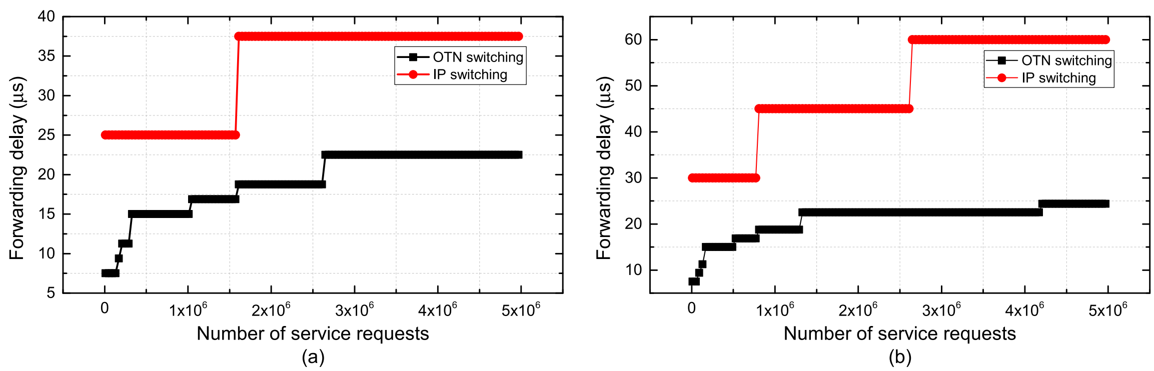


FIGURE 9. Average forwarding delay under different number of service requests: (a) Switching granularity is 1.25 Gb/s; (b) Switching granularity is 2.5 Gb/s.

that the BDT-OPC is relatively lower when port rate is 2.5Gbps/10Gbps, followed by the cases when the port rate is 40Gbps/100Gbps, 100Gbps/100Gbps, and 10Gbps/40Gbps under different switching granularity. In Figure, when the number of service requests is small, the BDT-OPC value exists a slight fluctuation. And when the number of service requests is greater than 0.3×10^6, the BDT based on optical port cost gradually remains stable. Therefore, we change the horizontal axis to the cost ratio of IP optical port cost is greater than the OTN optical port cost (i.e. the cost ratio of 5 indicates that the cost of 100 IP switching wavelengths is the same with the cost of 105 OTN switching wavelengths) when the number of service requests is 6×10^5. We can see in Figure 8 that the BDT-OPC is 24%-47% of the switching granularity when the port rate is 2.5Gbps/10Gbps, i.e. when the bandwidth of the service request is higher than 602 Mbps, the OTN switching has more advantages than the IP switching on optical port cost, and the OTN equipment can replace the IP router to satisfy the service requests. By contrast, the BDT-OPC for the 40Gbps/100Gbps, 100Gbps/100Gbps and 10Gbps/40Gbps port rate are 37%-71%, 50%-95%, and 81%-97% respectively. It indicates that the increased number of OTN equipment and optical cables for ensuring network connectivity is in a controllable range.

C. FORWARDING DELAY

We investigate the forwarding delay in this subsection. Figure 9(a) compares the average forwarding delay of service request in networks when the switching granularity is 1.25Gb/s, and the results verify that the latency is reduced effectively with OTN switching. The same conclusion can be obtained when the switching granularity is 2.5Gb/s. This attributes to the fact that layer 1 and layer 2 data structures need to be removed layer by layer when using IP switching. Therefore, it is difficult to achieve a low forwarding delay based on the layer 3 IP forwarding and switching. In contrast, optical switching is based on physical layer or media layer to switch and forward the service requests, a single-node forwarding delay of $10\mu s$ or less can be achieved. Simulation results in Figure 9 indicate that although the OTN switching results in an increase in the number of forwarding hops due to the fixed switching granularity, the forwarding delay is still smaller than the IP switching.

V. CONCLUSION

In this paper, optimization on a specific network topology is not the goal here. Instead we wish to obtain the bandwidth adaptation range for the two switching modes when OTN equipment are employed for traffic access and switching.

We abstract the switching behavior of OTN equipment and IP routers to establish a unified network-level connection model, and then conduct an adaptability study on CAPEX and forwarding delay for OTN switching and IP switching under different traffic parameters and network parameters.

The simulation results reveal that the IP switching is too expensive and a better way to reduce the hardware cost is employing optical transmission equipment to complete service switching in geographically distributed inter-datacenter networks. We validate that when average bandwidth between multiple datacenters is higher than 600Mbps, OTN equipment has more advantages than IP routers on capital expenditure, especially when port rate of equipment is high. In addition, the simulation results also verify that the forwarding delay is reduced obviously with OTN switching.

REFERENCES

[1] *Cisco Visual Networking Index: Forecast and Methodology, 2016–2021*, Cisco, San Jose, CA, USA, Sep. 2017. [Online]. Available: https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/complete-white-paper-c11-481360.html

[2] S. Jain *et al.*, "B4: Experience with a globally-deployed software defined WAN," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 43, no. 4, pp. 3–14, 2013.

[3] A. Asensio, M. Ruiz, and L. Velasco, "Requirements to support cloud, video and 5G services on the telecom cloud," in *Proc. Eur. Conf. Netw. Opt. Commun.*, Lisbon, Portugal, 2016, pp. 64–69.

[4] N. Bitar, S. Gringeri, and T. J. Xia, "Technologies and protocols for data center and cloud networking," *IEEE Commun. Mag.*, vol. 51, no. 9, pp. 24–31, Sep. 2013.

[5] O. Gerstel *et al.*, "Multi-layer capacity planning for IP-optical networks," *IEEE Commun. Mag.*, vol. 52, no. 1, pp. 44–51, Jan. 2014.

[6] A. Nikolaidis, "Leveraging FlexGrid and advanced modulations in a multilayer, inter-datacenter network," in *Proc. Opt. Fiber Commun. Conf. Exhib.*, Los Angeles, CA, USA, 2017, pp. 1–3, Paper W3D.1.

[7] S. Sengupta, V. Kumar, and D. Saha, "Switched optical backbone for cost-effective scalable core IP networks," *IEEE Commun. Mag.*, vol. 41, no. 6, pp. 60–70, Jun. 2003.

[8] *Router Off-Load Strategies. When is Router off-Load an Attractive Option*, Innov. Observatory Ltd., 2011, pp. 4–12.

[9] V. Gkamas, K. Christodoulopoulos, and E. Varvarigos, "A joint multi-layer planning algorithm for IP over flexible optical networks," *J. Lightw. Technol.*, vol. 33, no. 14, pp. 2965–2977, Jul. 15, 2015.

[10] I. Katib and D. Medhi, "IP/MPLS-over-OTN-over-DWDM multilayer networks: An integrated three-layer capacity optimization model, a heuristic, and a study," *IEEE Trans. Netw. Service Manag.*, vol. 9, no. 3, pp. 240–253, Sep. 2012.

[11] X. Zhao, V. Vusirikala, B. Koley, T. Hofmeister, and V. Kamalov, "The prospect of inter-data-center optical networks," *IEEE Commun. Mag.*, vol. 51, no. 9, pp. 32–38, Sep. 2013.

[12] F. A. Ghonaim, T. E. Darcie, and S. Ganti, "Impact of SDN on optical router bypass," *IEEE/OSA J. Opt. Commun. Netw.*, vol. 10, no. 4, pp. 332–343, Apr. 2018.

[13] F. A. Ghonaim, T. E. Darcie, and S. Ganti, "Adaptive router bypass using feedback adjusted OTN," in *Proc. Int. Conf. Telecommun.*, 2015, pp. 123–127.

[14] Y. C. Huang, Y. Yoshida, S. Ibrahim, R. Takahashi, A. Hiramatsu, and K. Kitayama, "Bypassing route strategy for optical circuits in OPS-based data center networks," *IEEE Photon. J.*, vol. 8, no. 2, Apr. 2016, Art. no. 0601510.

[15] *Software-Defined Networking: The New Norm for Networks*, Open Netw. Found., ONF White Paper, 2012, pp. 2–6.

[16] P. Lu and Z. Zhu, "Data-oriented task scheduling in fixed- and flexible-grid multilayer inter-DC optical networks: A comparison study," *J. Lightw. Technol.*, vol. 35, no. 24, pp. 5335–5346, Dec. 15, 2017.

[17] P. Papanikolaou, K. Christodoulopoulos, and E. Varvarigos, "Joint multi-layer survivability techniques for IP-over-elastic-optical-networks," *IEEE/OSA J. Opt. Commun. Netw.*, vol. 9, no. 1, pp. A85–A98, Jan. 2017.

[18] H. Yang, J. Zhang, Y. Zhao, J. Han, Y. Lin, and Y. Lee, "SUDOI: Software defined networking for ubiquitous data center optical interconnection," *IEEE Commun. Mag.*, vol. 54, no. 2, pp. 86–95, Feb. 2016.

[19] S. Das, G. Parulkar, and N. Mckeown, "Rethinking IP core networks," *IEEE Commun. Mag.*, vol. 5, no. 12, pp. 1431–1442, Dec. 2013.

[20] Z. Liu, J. Zhang, L. Bai, and Y. Ji, "Joint jobs scheduling and routing for metro-scaled micro datacenters over elastic optical networks," in *Proc. Opt. Fiber Commun. Conf.*, San Diego, CA, USA, 2018, pp. 1–3, Paper M2E.3.

[21] X. Cao *et al.*, "Adaptive DC interconnection provisioning in distributed all-optical micro-datacenters using holistically SDN orchestration for dynamic access," *Photon. Netw. Commun.*, vol. 35, no. 2, pp. 129–140, Apr. 2018.

[22] Y. Tan *et al.*, "Cost comparison and adaptability analysis for OTN switching and IP switching," presented at the 2nd Global Opt. Inf. Netw. Conf., Beijing, China, May 2018.

[23] *Interfaces for the Optical Transport Network*, document G. 709, ITU-T Recommendation, 2012.

[24] J. L. Vizcaíno, P. M. Krummrich, T. Jiménez, V. López, and Y. Ye, "OTN switching for improved energy and spectral efficiency in WDM MLR networks," in *Proc. Opt. Fiber Commun. Conf.*, Anaheim, CA, USA, 2016, pp. 1–3, Paper M3K.3.

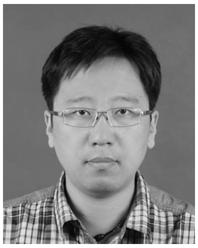
[25] C. Clos, "A study of non-blocking switching networks," *Bell Labs Tech. J.*, vol. 32, no. 2, pp. 406–424, 1953.

[26] C. Zhang, H. Yang, and G. F. Riley, "Admission control in software-defined datacenter network in view of flow table capacity," in *Proc. IEEE Conf. Comput. Commun. Workshops (INFOCOM)*, Honolulu, HI, USA, Apr. 2018, pp. 871–876.

YANXIA TAN received the B.S. degree from the Chongqing University of Posts and Telecommunications in 2013. She is currently pursuing the Ph.D. degree with the State Key Laboratory of Information Photonics and Optical Communications, Beijing University of Posts and Telecommunications. Her research interest focused on flexible optical network, IP-optical integration network design, and optimization.

RENTAO GU received the B.E. and Ph.D. degrees from BUPT, Beijing, China, in 2005 and 2010, respectively. From 2008 to 2009, he was a Visiting Scholar with the Georgia Institute of Technology, Atlanta, GA, USA. He is currently an Associate Professor with the School of Information and Telecommunication Engineering, BUPT. His current research interests include access network and intelligent information processing.

YUEFENG JI received the Ph.D. from BUPT, Beijing, China. He is currently a Professor and the Executive Deputy Director of the State Key Laboratory of Information Photonics and Optical Communications, BUPT. His research interests are primarily in the area of broadband communication networks and optical communications, with emphasis on key theory, realization of technology, and application. He is a fellow of CIC, CIE, and IET.



DAJIANG WANG received the master's degree from Shanghai University, Shanghai, China. He is currently a Senior Engineer with ZTE Corporation. In recent years, he has mainly researched on AI for OTN, software defined optical network, and synergy of intelligent controlling and management for OTN.

HONGBIAO LI received the master's degree from Hohai University, Nanjing, China. He is currently a Chief Engineer with ZTE Corporation. Meanwhile, his research and work have been focused on the planning and design about IP-optical integration solution for many years.