SDN Enabled Restoration With Triggered Precomputation in Elastic Optical Inter-Datacenter Networks

Yu Xiong, Yuanyuan Li, Bin Zhou, Ruyan Wang, and George N. Rouskas

***Abstract*—Flexi-grid elastic optical networks (EONs) provide a promising infrastructure for meeting the high-bandwidth and low-latency requirements of interconnecting datacenters. Since failures of high-capacity fiber links prevent users from accessing cloud services and lead to large amounts of data loss, much recent research has focused on inter-datacenter network failure recovery. In this paper, we build upon software-defined networking (SDN) technology and develop a fast restoration strategy with triggered precomputation (FR-TP) that achieves fast failure recovery with minimal resource overhead. Our work makes several contributions. We extend the controller functionality and OpenFlow protocol for software-defined elastic optical inter-datacenter network architectures, so as to obtain the global network state information quickly and accurately. We then present the FR-TP strategy and an associated trigger mechanism to compute backup paths before a link failure occurs. Furthermore, to improve the efficiency of path computation and bandwidth resource assignment, we construct a layered auxiliary graph of spectrum window planes (SWP-LAG) using a residual capacity matrix to change dynamically the width of the spectrum window planes (SWPs) to satisfy different service requests. Simulation results demonstrate that, compared with existing restoration strategies, the proposed FR-TP strategy combined with SWP-LAG reduces the recovery time by up to 30.4% in the network topologies studied without increasing the blocking probability.**

***Index Terms*—Elastic optical networks (EONs); Inter-datacenter; Layered auxiliary graph; Restoration; Software defined networking; Triggered precomputation.**

I. INTRODUCTION

With the rapid development of cloud services, such as cloud computing, big data, and high definition video streaming, the datacenter, as an important carrier providing the capacity of information storage and processing, has become the core infrastructure to support the next generation of Internet technology. Meanwhile, the vast amounts of datacenter services interacting among multi-datacenters (multi-DCs) often show high-burstiness and high-bandwidth characteristics. In addition, the information flow among datacenters grows continually, so that the demand for high-performance communication among massive datacenters becomes more prominent [1,2]. However, the traditional interconnected datacenter network in which electrical switching is the core technology meets technical bottlenecks in terms of bandwidth capacity, energy consumption, and transmission distance. Therefore, optical interconnected-datacenters (inter-DCs) provides one of the key technologies to satisfy the requirements of large-scale datacenter networking with the advantages of large capacity, high bandwidth, and low latency.

As shown in [3], datacenter networks interconnected by high-speed optical networks are growing at an immense rate. For the diversity and point-to-content characteristics of datacenter services, they urgently need optical network transport to provide flexibility, speed, and high-level quality of service (QoS) [4]. The emerging flexible-grid elastic optical networks utilize bandwidth variable transponders (BV-Ts) and wavelength selective switches (BV-WSSs) that operate on a series of spectrally contiguous frequency slots (FSs) to set up lightpaths. What is more, EONs can provide optical connectivity for a large variety of bandwidth requests ranging from 1 Gbps to 1 Tbps, so EONs have evolved as next-generation optical networks with finer granularity (e.g., 12.5 or 6.25 GHz) [5,6]. Since these FSs have much narrower bandwidth than conventional wavelength channels, EONs can provision bandwidth adaptively according to actual traffic demands. Hence, the EON is a crucial physical infrastructure to enable communication between large-scale servers and forms the foundation of network storage and computation [7].

However, as the main carrier of data services, the capacity of a single fiber in EONs has been improved to Pb/s [8]. Therefore, any one link failure in inter-datacenter networks interconnected by EONs will result in a huge amount of data loss. To this end, it is crucial to study and design efficient network control architecture and failure recovery mechanisms for elastic optical inter-datacenter networks. Recently, as a promising centralized control architecture, software-defined networking (SDN) based on OpenFlow (OF) protocol has gained popularity by supporting programmability of datacenter and network functionalities [9,10], which can provide maximum flexibility for the

Manuscript received June 5, 2017; revised October 25, 2017; accepted October 27, 2017; published December 27, 2017 (Doc. ID 296669).

Y. Xiong (e-mail: cqxiongyu@foxmail.com), Y. Li, B. Zhou, and R. Wang are with the Key Laboratory of Optical Communication and Networks, Chongqing University of Posts and Telecommunications, Chongqing, China.

Y. Xiong and G. N. Rouskas are with the Department of Computer Science, North Carolina State University, Raleigh, North Carolina 27695, USA.

G. N. Rouskas is also with the Department of Computer Science, King Abdulaziz University, Jeddah, Saudi Arabia.

https://doi.org/10.1364/JOCN.10.000024

operators, make a unified control over various resources, and abstract them as a unified interface for the joint optimization of functions and services with a global view [11,12]. Compared with generalized multiprotocol label switching (GMPLS), SDN centralizes intelligent control functions into a controller by abstracting a control plane away from the data plane. It can enable flexible control of network traffic and provides an innovative platform for network applications. A number of studies have been carried out in recent years, and OF extensions have been developed to implement elastic optical networks in SDN architecture, demonstrating the extensibility of the SDN framework [13,14]. However, only a few of them deal with network reliability (e.g., protection and restoration) since OF was originally designed for supporting packet switching only [15]. Therefore, if a proper recovery mechanism can be implemented into the SDN architecture, a more reliable network would be realized.

To this end, we propose a novel fast restoration strategy with triggered precomputation (FR-TP) for software-defined elastic optical inter-datacenter networks in this paper. The strategy can achieve intelligent precomputation failure management by extending the SDN architecture that orients elastic optical inter-datacenter networks. Our theoretical analysis and simulation results indicate that the proposed FR-TP can provide instant recovery time, low blocking probability (BP), and high resource utilization. The main contributions of this work are as follows.

- To the best of our knowledge, this is the first work to provide the triggered precomputation of a restoration path before a link failure occurs in software-defined elastic optical inter-datacenter networks.
- We redesign the software-defined elastic optical inter-datacenter network architecture by extending OF protocol and functionality of the controller for datacenter application requests.
- We establish a SWP-LAG that can dynamically change the width of SWP through a residual capacity matrix to satisfy different sizes of service requests. Meanwhile, we design a distance-adaptive one-step dynamic routing and spectrum allocation (DRSA) scheme based on SWP-LAG.
- Under the background of software-defined EONs, we provide detailed analysis and comparison of FR-TP and other strategies in terms of BP, resource occupation rate, failure restoration ratio (FRR), and recovery time. Then we prove that FR-TP can not only achieve extremely fast recovery, but also guarantee the actual availability of the required resources after link failure occurs.
- We perform and simulate a software-defined elastic optical inter-datacenter network testbed, which we built with Mininet + Floodlight and C++ tools.

The rest of this paper is organized as follows. Section II summarizes related work. In Section III, we design the software-defined elastic optical inter-datacenter network architecture and extend the OF controller to support the proposed strategy. In Section IV, the FR-TP strategy is proposed for datacenter services under this control architecture. Section V introduces a distance-adaptive DRSA (DA-DRSA) heuristic algorithm based on the SWP-LAG model. Numeric results and analysis are given in Section VI. Finally, Section VII concludes the paper and presents some future work.

II. RELATED WORK

According to whether the backup resource is reserved for link failure, the failure management for elastic optical networks could be divided into two categories: protection and restoration. Compared with protection strategies, restoration strategies have higher resource utilization rates and lower BP because the restoration need not reserve backup resources before the failure occurs. According to the restoration path calculated before or after the failure occurs, we further separate the restoration technologies into reactive restoration and proactive restoration. Reactive restoration calculates a restoration path after the failure occurs, but proactive restoration calculates a restoration path in the way of precomputation before the failure occurs. So far, several studies have been concerned with reactive restoration strategies for failure services in software-defined EONs. In [16], a novel multi-stratum resource resilience (MSRR) architecture is proposed for datacenter services in software-defined datacenter interconnection based on IP over EON, which effectively improves resource utilization. However, due to the recovery routing algorithm performed in the IP layer, it leads to high delay of failure recovery. Reference [17] presents OF-enabled dynamic restoration in EONs, detailing the restoration framework and algorithm, OF protocol extensions, and analyzing the restoration performance via experimental validation. Reference [18] presents a proof-of-concept demonstration of OF-controlled dynamic recovery for only one failed connection on a real EON testbed employing a flexible transmitter (Tx) and receiver (Rx). The authors of [19] proposed a fast OF-based restoration scheme to minimize the parallel hardware configuration delay. However, physical impairments in the network are not considered. Reference [20] investigates OF-based dynamic lightpath restoration in EONs by considering physical layer impairments, modulation formats, or even bandwidth squeezed restoration. In [21], a novel reactive restoration strategy called SDN-ind is proposed. This reactive restoration strategy skillfully leverages the controller to trigger an independent OF communication after each backup path computation. It makes each optical switch traversed by the backup path configured in parallel, which greatly improves the recovery time. However, these strategies do not consider the problem that the calculation time of the rerouting accounts for a larger proportion of the total recovery time. In [22], a proactive restoration strategy is addressed to reduce the calculation time of the rerouting procedure during the restoration scheme. However, in the distributed network based on GMPLS control technology, the precomputation is achieved through interworking with a large number of control signals among the nodes, and the transmission of these signals will consume a large amount of network resources. In elastic optical inter-datacenter networks, the proactive restoration strategy precomputes the restoration path by leveraging SDN technology, occupying only a small part of the computing resources of the OF controller. In addition, the SDN controller can quickly and accurately converge the network state information with a global view and configure the network nodes in parallel. The main feature of SDN in support of restoration is its inherent capability of configuring several elements (e.g., BV-Ts and BV-WSSs) in parallel, thus reducing the total configuration time. Thus, the performance for a proactive restoration strategy based on SDN control technology may be superior.

In light of the above, there is a growing requirement to address the trade-off problem between recovery time and resource overhead. However, the solution of this issue is closely related to routing and spectrum allocation (RSA). Especially during the restoration process, RSA is an important factor that directly affects the restoration success probability, resource utilization, and recovery time. In addition, the spectrum allocation in EONs needs to ensure two constraints that increase the difficulty of RSA. The two constraints are as follows: (1) the spectrum continuity constraint, i.e., the optical path must use the same spectrum block for the optical network without spectrum conversion capability; and (2) the spectrum contiguity constraint, which requires the allocated spectrum to be chosen from contiguous subcarrier frequency slots in the frequency domain on each link. Therefore, it is necessary to design an efficient RSA to optimize network resources. In [23], a dynamic multi-path provisioning algorithm for EONs is proposed. The algorithm tries to set up dynamic connections with multi-path provisioning, which reduces the BP. However, this algorithm may over-reserve bandwidth. In [24,25], a distance-adaptive spectrum allocation algorithm is proposed, which is proved to use the least amount of resources and attain high spectrum efficiency. Moreover, in [26] the RSA problem is divided into two phases: a routing and modulation level phase and a spectrum allocation phase, and they are solved one by one. However, when they finish computation of the path, they cannot guarantee whether the spectrum slots on the calculated path are available and whether the selected modulation from 24QAM to BPSK is optimal. From the discussion above, we can find that most of the previous studies are targeted for two-step RSAs, which can satisfy neither the spectrum availability nor the two constraints of EONs when implementing routing. To solve these issues, in this work, we provide a promising dynamic RSA of datacenter services by designing an efficient distance-adaptive one-step RSA scheme based on SWP-LAG. This RSA scheme not only satisfies the two constraints, it also guarantees the availability of spectrum.

Our previous work in [27] studies the precomputation-based restoration path (P-RP). This paper extends the work in [27] with a restoration-strategy-based triggered precomputation (FR-TP), which performs a proactive restoration strategy that considers the availability guarantee of recovery resources. Moreover, this paper is still considered a promising solution for the DRSA of datacenter services by designing an efficient distance-adaptive one-step DRSA scheme based on SWP-LAG. We design a novel residual capacity matrix for SWP-LAG to dynamically change the width of SWPs to satisfy different sizes of service requests.

III. Network Model

The SDN-based elastic optical inter-datacenter network architecture is illustrated in Fig. 1. The network is divided into the control plane and the underlying data plane. The underlying elastic optical inter-datacenter network is centrally managed by the controller with extended OF protocol. To realize the control functionality of the controller, the network nodes on the underlying data plane should be composed of an OF protocol agent and an OF-enabled bandwidth variable optical switch (OF-BVOS). The OF protocol agent receives flow table messages from the controller, and then translates them into the logical language, which the underlying hardware devices can understand, and then controls the cross-connection process of the underlying OF-BVOS. In the case of link failures, the failure detection module will discover them and deliver the failure information to the failure restoration module, which decides to apply a restoration strategy associated with the optical network resources. Finally, configuration commands are issued to complete the restoration in parallel by the controller through the extended OF protocol.

To support the functionalities of the above architecture, the controller and OF protocol need to be extended. The datacenter ID is added to the OF protocol and the functionality module of the controller is extended, as shown in Figs. 2 and 3.

A. Functional Module of SDN Controller

The OF controller consists of seven modules, i.e., network abstraction, path computing entity (PCE) and plug-in, data base module, forward Open API, precomputation and restoration, failure detection, and Web management. The responsibilities and interactions among the functional modules are provided below. The network abstraction

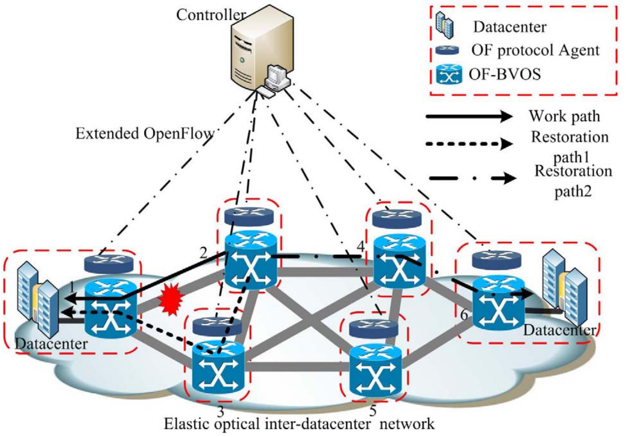


Fig. 1. Architecture of software-defined elastic optical inter-datacenter network.

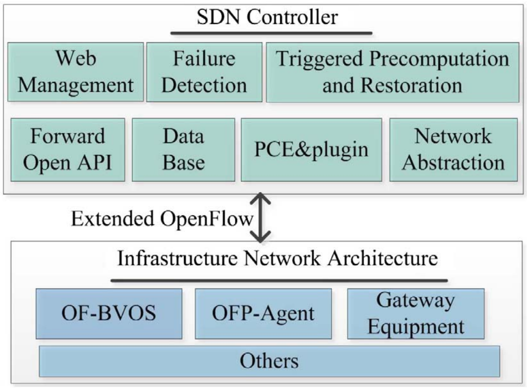


Fig. 2. Functional module of SDN controller in software-defined elastic optical inter-datacenter networks.

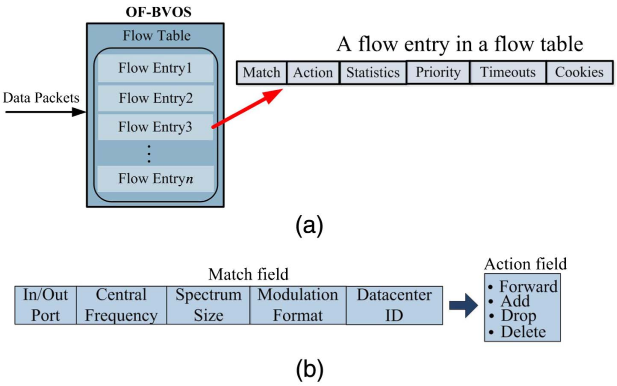


Fig. 3. Extensions of OF protocol. (a) The flow table in OF-BVOS. (b) The extensions of flow entry.

module can abstract the required flexible optical resources, while the failure detection module interworks the information with OF-BVOS periodically to perceive optical networks through the extended OF protocol. In case of link failures, the failure detection module discovers them and delivers such failure information to the triggered precomputation and restoration module. Before the link failure occurs in an EON, the precomputation and restoration module decides to apply a precomputation strategy associated with the optical network resources. If the precomputed resources are occupied by newly arrived services, the controller will immediately enable the precomputation module again to ensure the availability of the required resources. The PCE module can calculate the working and restoration lightpaths interworking with the OF protocol agent of the OF-BVOS, where the various computation strategies are alternative as a plug-in. The information of paths are conserved into data base management, and the results of paths are updated.

B. OF Extensions

In each OF-BVOS, a flow table, working as the decision-making basis, is set up by a series of rules. Each flow table contains a set of flow entries, which are defined as "Match," "Action," "Statistics," "Priority," "Timeouts," and "Cookies," as shown in Fig. 3(a). An arriving request is compared to each flow entry in the flow table; if it matches with a certain flow entry, the exact instruction of that entry will be executed, otherwise a new entry will be created or the service request will be dropped. In this way, the controller gets remote access to handling the service requests. To satisfy the requirement of elastic optical inter-datacenter networks, we extend the FLOW-MOD message to flow entry for the current OF protocol, i.e., OF v1.3.

The extension of the match field in flow entry is shown in Fig. 3(b). The fields "In/Out Port," "Central Frequency," "Spectrum Size," "Modulation Format," and "Datacenter ID" are extended to the "Match" part of flow entry as the main features in elastic optical inter-datacenter networks. The datacenter ID is added to mark different datacenters and facilitates data center service addressing. The "Action" part of flow entry, including "Forward," "Add," "Drop," and "Delete," are used to set up or release a lightpath. By means of matching the FLOW-MOD message to the flow entries, the OF-BVOS can execute the exact instruction that is ordered by the controller.

IV. FAST-RESTORATION-BASED TRIGGERED PRECOMPUTATION

The precomputation restoration strategy is proposed to reduce the high weight of calculation time of the rerouting procedure during the restoration scheme [22]. However, in this distributed network, precomputation is achieved through interworking with a large number of control signals among the nodes (such as GMPLS), while the transmission of these signals will consume a large amount of network resources.

A. Triggered Precomputation Before a Failure Occurs

To effectively reduce the recovery time within a centralized control plane in EONs, we leverage SDN technology to precompute the restoration path based on the model established in the above section, which occupied just a small amount CPU computing resources through an extra opening thread in the controller. Precomputation of a restoration path can be achieved without consuming the transmission resources between nodes. On the other hand, the SDN controller can quickly and accurately converge the network state information with a global view and make the network nodes configured in parallel. The precomputation restoration strategy, first, precomputes the restoration path for the delay-sensitive and high-level services associated with the optical network resources before the interruption. At the same time, the path information is stored in the controller. Then, when service is really interrupted, the controller can directly use the precomputation recovery strategy that is stored in the controller and allocate the network resources for the interruption of service. Thus, the restoration route is successfully set up.

The resources of the precomputed path are calculated and stored in the controller before the failure occurs, but no reservation of the "backup resources" is performed until failure occurs. However, on the one hand, idle resources on the fiber links face competition under heavy load for the massive bandwidth requirements of the arriving requests. The resources on the precomputed restoration path may be occupied at any time. On the other hand, when disasters occur, such as earthquake, tsunami disasters, and human damage, there may be a large number of precomputed fiber link failures at the same time. Furthermore, in general, the failure occurrence time cannot be predicted. Therefore, there is no guarantee that the precomputed resources on the restoration path are still available under heavy load or when the failure occurs, so that the failure recovery ratio cannot be guaranteed. To ensure that the network has a high restoration success ratio, we design a simple but effective FR-TP restoration strategy to ensure the availability of resources on the restoration path. The implementation procedure of FR-TP restoration strategy is as follows.

First, we precompute the restoration path for the delay-sensitive and high-level services associated with the optical network resources before the interruption. Then, we set up triggered alarm information based on the availability of the resources on the precomputed restoration path. When the resources on the precomputed restoration path are occupied by other newly arriving services or the precomputed paths also fault, the alarm information is triggered and causes the controller to carry out the precomputation. At this time, the controller is required to recalculate a restoration route with available resources according to the updated network state. If the resources on the precomputed restoration route have not been occupied all along by other services, the process will not trigger the alarm information. Finally, when the working path is really interrupted, the controller can directly use the precomputation recovery strategy that is stored in the controller and provision the network resources for the interruption of service. The FR-TP strategy avoids path computation time during the recovery process and guarantees the availability of required resources. Thus, it can effectively reduce the recovery time and ensure high recovery success probability.

B. Restoration After a Failure Occurs

When link failures occur in the network, each of them is detected by the closest OF-BVOS and the information is sent to the controller by PORT_STATUS messages. The controller searches for the precomputation restoration route information stored in the controller. Then, according to the route information, the controller sends FLOW_MOD messages labeled with "ADD" and "DELETE" to the source node. These messages should cover all the affected paths, which means all the flow entries of the affected paths are deleted in the source node. Thus, only the available flow entries remain in the flow table of the source node, and the packet will be forwarded along the remaining paths. A few more FLOW_MOD messages are sent to delete the rest of the useless flow entries in the OF-BVOSs of all the affected paths. Thus far, a restoration based on precomputation is accomplished in the software-defined elastic optical inter-datacenter network. This is a break-before-make strategy. The interworking procedure for restoration based on precomputation, when link failure occurs, is shown in Fig. 4, where τ_{ps} is the time used to upload the failure information from the failure link node to the controller, τ_{search} is the time used to query the precomputation routing strategy stored in the controller, and τ_i represents the time used by the controller to send FLOW-MOD messages labeled with ADD to the nodes of the precomputed paths. The nodes include the source node, intermediate nodes, and datacenter node (i.e., destination node). τ_i' represents the time used by the controller to send FLOW_MOD messages labeled with "DELETE" to the nodes of the affected paths, and Δt is the time interval between when the controller issues the message labeled with "ADD" and the message labeled with "DELETE". In general, $\Delta t \geq 0$.

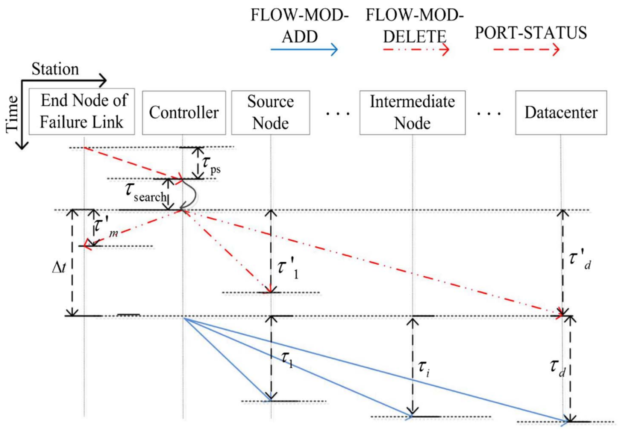


Fig. 4. Precomputation restoration strategy based on SDN timeline.

Hence, the restoration time of the whole restoration strategy based on precomputation is

$$\mathrm{t_{pc}} = \tau_{\mathrm{ps}} + \tau_{\mathrm{search}} + \mathrm{Max}\left[\left(\Delta t + \underset{i\in[1,n]}{\mathrm{Max}}\,\tau_i'\right), \underset{i\in[1,n]}{\mathrm{Max}}\,\tau_i\right]. \qquad (1)$$

The traditional restoration strategy has larger delay, which is produced mainly by the process of computing bandwidth, center frequency, and the cross-connection process between OF-BVOSs. In addition, the procedure of computing bandwidth and center frequency account for a large proportion of the restoration time. The restoration strategy based on precomputation can solve the problem of large delay produced by bandwidth and center frequency computation.

V. DISTANCE ADAPTIVE DRSA BASED ON THE SWP-LAG MODEL

One of the key technologies ensuring high bandwidth transmission in EONs is RSA. Different from the routing and wavelength assignment in WDM optical networks, the fine granularity of EONs poses new challenges, for instance, how to satisfy the spectrum continuity and spectrum contiguity constraints, how to decrease the influence of the spectrum fragmentation rate, and how to improve the resource utilization rate. To address these issues, we design a new distance-adaptive dynamic RSA with the SWP-LAG model to improve the success probability and efficiency of RSA.

A. Layered Auxiliary Graph of Spectrum Window Plane

Under spectrum continuity and contiguity constraints in EONs, to find the route with continuous spectrum resources, we introduce the Spectrum Window (SW) concept [28] to implement the one-step RSA algorithm. In the example of Fig. 5, assume that a fiber has N FSs, and since each SW is made up of one FS, there are N SWs of a fiber. Then, based on the concept of SW, another concept called spectrum window plane (SWP) is defined, which is similar to the concept of a waveplane [29] in the traditional WDM network. In a WDM network, each waveplane corresponds to a specific wavelength. Likewise, each SWP corresponds to a SW in an EON. Then, we leverage SWP and the residual capacity matrix of spectrum slots to build a layered auxiliary graph (LAG) model. For example, the LAG model of a network topology with 6 nodes and 10 links is shown in Fig. 6. Each SWP corresponds to a residual capacity matrix R_N, which represents the usage of spectrum slot resources on each link in the Nth SWP:

$$R_N = \begin{array}{c c} & \begin{array}{cccc} v_1 & v_2 & \cdots & v_n \end{array} \\ \begin{array}{c} v_1 \\ v_2 \\ \vdots \\ v_n \end{array} & \left[\begin{array}{cccc} - & m_{12} & \cdots & m_{1n} \\ & - & \cdots & \vdots \\ & & - & m_{n-1,n} \\ & & & - \end{array}\right] \end{array}. \quad (2)$$

Here, N is a positive integer. The element 1 indicates that link (i,j) has available spectrum resources; if the value is 0 that indicates that there are no available spectrum resources on link (i,j).

Because of the diversity of datacenter services in the network, different services often need to allocate available spectrum resources with different bandwidth sizes. Therefore, the multi-spectrum slot window plane (MSWP) is built based on SWP. The calculation method of the corresponding residual capacity matrix of SW with a multi-spectrum slot is

$$R(\xi, n_i) = \bigcap_{N \in (\xi, \xi+n_i)} R_N, \quad (3)$$

where ξ represents the number of starting FSs and n_i represents the number of required FSs for the services. Finally, we can map out the LAG of the multi-spectrum slot width according to the residual capacity matrix.

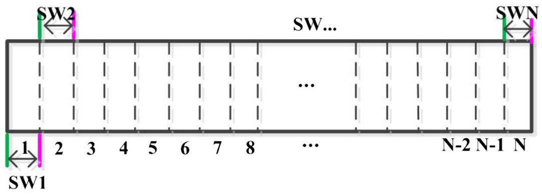


Fig. 5. SWs in a fiber link.

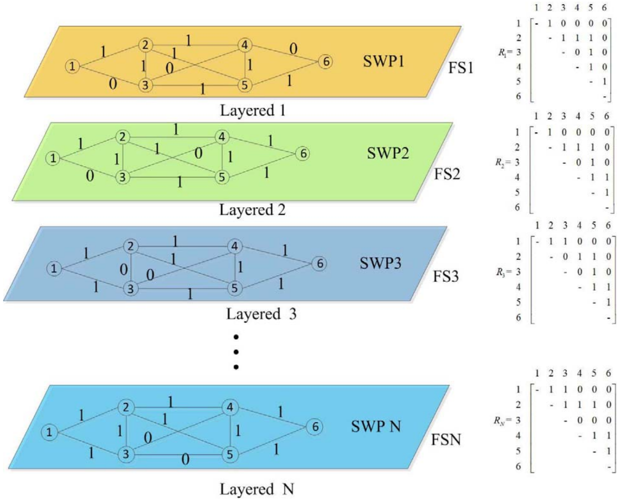


Fig. 6. Example for LAG of SWP.

B. Distance-Adaptive Dynamic RSA

The signal of an EON can be transported by different modulation formats according to the actual transmission distance of the service. In transparent optical networks, with each bit increasing in the modulation format, the transmission distance must be halved to keep the same level of quality of transmission (QoT) [30], as shown in Table I. Hence, the modulation format should be lowered as the transmission distance of the optical path increases in order to resist against deterioration. The bandwidth of each FS is assumed to be 12.5 GHz, and each row shows the FS capacity and transparent reach of each modulation format. We consider only four modulation formats, which are BPSK, QPSK, 8QAM, and 16QAM.

While most requests are anycast requests in inter-datacenter networks, the destination node of such a request is not fixed, and any datacenter with the required

TABLE I
PARAMETERS OF MODULATION FORMATS [28]

Modulation Format	Subcarrier Capacity (Gbit/s)	Transparent Reach (km)
BPSK	12.5	4000
QPSK	25	2000
8QAM	37.5	1000
16QAM	50	500

service content could be selected as the destination node. Here, D is defined as the set of datacenters, and $|D|$ denotes the total number of datacenters. In this study, the same application services are stored in all datacenters. A distance-adaptive modulation scheme based on the Dijkstra algorithm is used when establishing a connection request [24]. Any service r can find $|D|$ shortest paths from the source node to each datacenter with the Dijkstra algorithm. For a request $r(s, d, \Phi)$, where s, d, and Φ are the source node, destination node, and bandwidth of service requests in Gbit/s, respectively, the required number of FSs for a certain modulation format can be easily computed by

$$M_i = \mathrm{Mod}\left(\sum_e L_e\right), \qquad e \in R_{s,d,i}, \tag{4}$$

$$n_\mathrm{i} = \frac{\Phi}{B \cdot \log_2 M_\mathrm{i}}, \tag{5}$$

where $\mathrm{Mod}(\cdot)$ returns the highest modulation level that transmission distances can support, L_e is the length of each link, n_i is the required FSs, B denotes the size of a FS when the modulation is BPSK ($B = 12.5$ GHz), and M_i is the modulation level being used.

Then, we design a DA-DRSA heuristic algorithm based on the SWP-LAG model. We use two variables, *path* and *index*, to record the information of the working route and starting FS index of the working path, respectively. The details of the DA-DRSA algorithm are introduced in Algorithm 1.

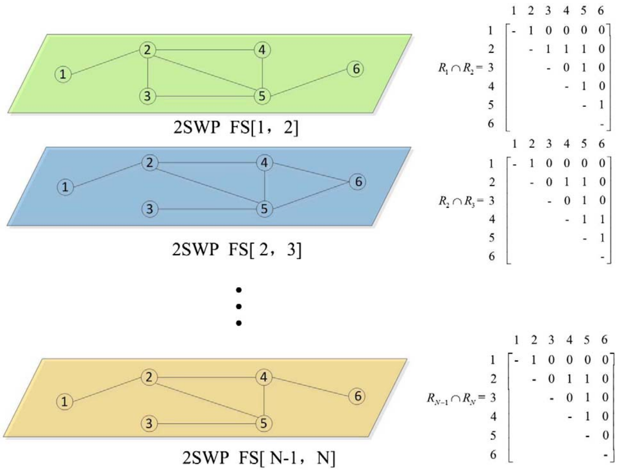
Algorithm 1: DA-DRSA Algorithm

```
Input: Network topology G(V, E, D) a set of requests $R = \{r_1, r_2, r_3 \ldots\}$, r = r$(s, d, \Phi)$.
Output: path; index, d;
for each modulation formation (from 16QAM to BPSK in Table I) do
    Calculate the number of required FSs according to Eqs. (4) and (5), namely, the width of SWP;
    Calculate the residual capacity matrix according to Eq. (2) for each MSWP;
    Map out the LAG of MSWP according to the matrix;
    for each MSWP (from the lowest to highest index) do
        Use Dijkstra's algorithm to find a shortest route $p_{wi}$ from s to each datacenter d, the total number is $|D|$;
        Sort the $\{p_{wi}, i = 1, 2, \ldots, |D|\}$ based on C in ascending order;
        for each $\{p_{wi}\}$ (from the lowest to the highest index) do
            if $p_{wi}$ is found and the transmission distance of $p_{wi}$ is shorter than the transparent reach of the current $M_\mathrm{i}$ then
                if path $= NULL$ then
                    $p_{wi} \rightarrow$ path, start index of current MSWP $\rightarrow$ index;
                else
                    if the hop length of $p_{wi}$ is smaller than that of route then
                        $p_{wi} \rightarrow$ path, start index of current MSWP $\rightarrow$ index;
                    end if
                end if
            else
                Move to next MSWP;
            end if
        end for
    end for
end for
if path $= NULL$ then
    Move to next modulation;
end if
return (path, index, d)
```

Figure 7 illustrates an intuitive example of the DA-DRSA algorithm with SWP-LAG. The service request is $r = (1, 6, 4)$, which needs four FSs from source node 1 to destination node 6. First, we adaptively select the QPSK modulation format according to the transmission, and the number of required spectra is calculated as two according to Eq. (5). Then, we perform RSA in the network topology of 6-nodes 10-links (n6s10) in Fig. 6. We calculate the residual capacity matrix for each 2SWP(M=2) from SWP1 with the lowest index. If the element in the matrix is 1, it shows that there is a link between the corresponding nodes. If the element is 0, it shows that there are no available continuous spectrum slots between the corresponding nodes and this link should be deleted. Finally, we map out the LAG of 2SWP according to the matrix, as shown in Fig. 7. Thus, we can find a routing 1–2–5–6 for r occupying FS [1,2] in the first 2SWP. The one-step RSA algorithm can be implemented directly on each 2SWP in the LAG.

VI. SIMULATION ANALYSIS

In this section, to demonstrate that the FR-TP restoration strategy can meet network survivability requirements, we leverage the Mininet + Floodlight and C++ simulation



$$R_1 \cap R_2 = \begin{array}{c|cccccc} & 1 & 2 & 3 & 4 & 5 & 6 \\ \hline 1 & - & 1 & 0 & 0 & 0 & 0 \\ 2 & & - & 1 & 1 & 1 & 0 \\ 3 & & & - & 0 & 1 & 0 \\ 4 & & & & - & 1 & 0 \\ 5 & & & & & - & 1 \\ 6 & & & & & & - \end{array}$$

$$R_2 \cap R_3 = \begin{array}{c|cccccc} & 1 & 2 & 3 & 4 & 5 & 6 \\ \hline 1 & - & 1 & 0 & 0 & 0 & 0 \\ 2 & & - & 0 & 1 & 1 & 0 \\ 3 & & & - & 0 & 1 & 0 \\ 4 & & & & - & 1 & 1 \\ 5 & & & & & - & 1 \\ 6 & & & & & & - \end{array}$$

$$R_{N-1} \cap R_N = \begin{array}{c|cccccc} & 1 & 2 & 3 & 4 & 5 & 6 \\ \hline 1 & - & 1 & 0 & 0 & 0 & 0 \\ 2 & & - & 0 & 1 & 1 & 0 \\ 3 & & & - & 0 & 1 & 0 \\ 4 & & & & - & 1 & 0 \\ 5 & & & & & - & 0 \\ 6 & & & & & & - \end{array}$$

Fig. 7. Example of the DA-DRSA algorithm based on SWP-LAG.

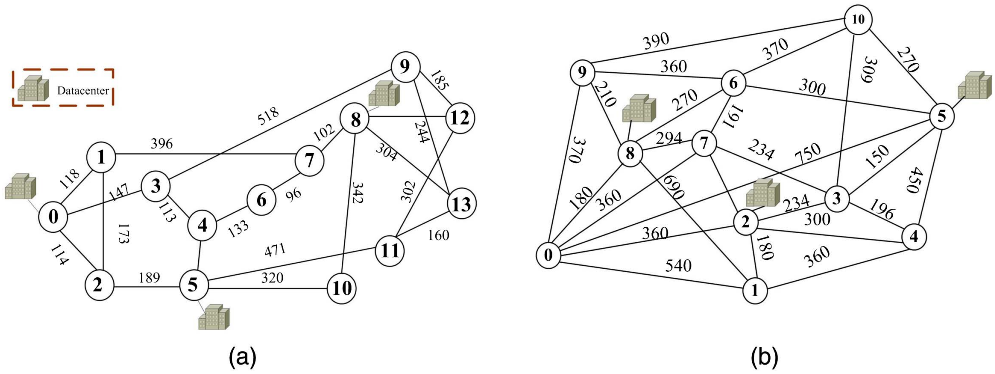
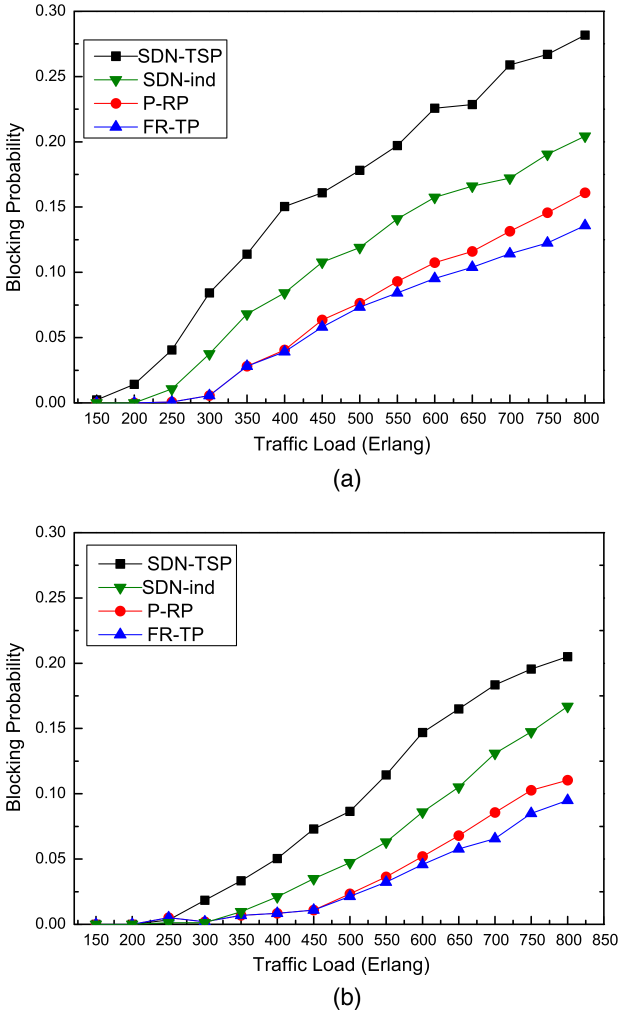


Fig. 8. Network topology: (a) NSFNet and (b) COST239.

tools to build a test platform in the network topology of NSFNet (14 nodes, 21 links, average node degree is 3) and the COST239 network (11 nodes, 26 links, average node degree is 4.73) as shown in Fig. 8. NSFNet has datacenters at nodes 0, 5, and 8, and the COST239 network has datacenters at nodes 2, 5, and 8 [31]. We evaluate the performance of the FR-TP strategy under a heavy traffic load scenario and compare it with the traditional shared-protection-based SDN (SDN-TSP), SDN-ind, and P-RP in [27]. SDN-TSP is a shared protection strategy sharing multiple backup resources for multiple working primary paths without bandwidth squeezing, and SDN-TSP has the same datacenters as the FR-TP network model. SDN-ind is from [21], but only a single-link failure scenario is considered here. Both SDN-TSP and SDN-ind use a general two-step RSA algorithm. Assume that there is one pair of bi-directional fiber on each link, and the available spectrum width of each fiber is set to 4000 GHz with width of 12.5 GHz. In this simulation, the flow requests to datacenter nodes are established with bandwidth randomly distributed from 12.5 to 100 Gbps, and they arrive at the network following a Poisson process with the arrival rate λ. The service holding time follows a negative exponential distribution with departure rate μ, and the traffic load is λ/μ (Erlang). The single node configuration time is about 50 ms, and the processing time for packets is about 2 ms. These time parameter settings are the same as those of Ref. [21].

In this paper, three performance indices are considered, which are BP, FRR, and average recovery time (ART). BP is the ratio of the number of services rejected by the network over the number of all the arrival services in the network. For each failure, FRR is defined as the ratio between the number of successfully recovered services and the total number of services affected by the failure; ART refers to the link failure recovery time, which consists of the failure detection time and provisioning time of a new path.

The BPs of different strategies for NSFNet and COST239 can be seen in Fig. 9. First, in both simulation scenarios, SDN-TSP has the highest BP among the four strategies due to the resource consumption of the protection path. In contrast, FR-TP has the lowest BP. That is because the protection algorithm needs to reserve protection resources on a backup path for each service in advance, which occupies many spectrum resources of the network. FR-TP and P-RP have lower BP than SDN-ind, because FR-TP and P-RP use a DA-DRSA scheme based on SWP-LAG, which can transmit the same services with less spectrum resources. However, the BPs of FR-TP and P-RP have almost the same performance, although FR-TP has relatively larger fluctuation than P-RP with increase of



(a)

(b)

Fig. 9. BP: (a) NSFNet and (b) COST239.

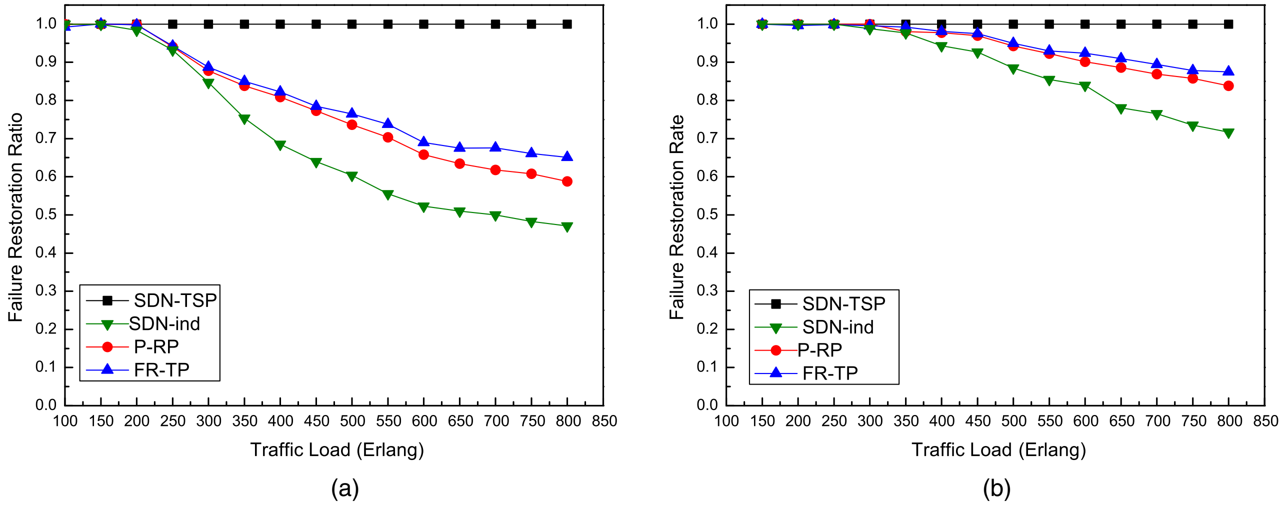


Fig. 10. FRR: (a) NSFNet and (b) COST239.

traffic load, i.e., FR-TP has better performance than P-RP under overload. This is because the FR-TP strategy can guarantee the availability of the required resources on the precomputed path in a timely manner, so it can avoid the blocking caused because no "backup resources" are reserved, especially under high traffic load.

In Fig. 10, the FRRs of different strategies for NSFNet [Fig. 10(a)] and COST239 [Fig. 10(b)] are presented. As shown in these figures, the average FRR of SDN-TSP is 100% due to the protection resources reservation on the backup path for each service, which can support the complete reliability assurance for each working path. Likewise, the FRRs of three restoration strategies in the COST239 network are higher than those in the NSFNet network because of the higher network node connectivity. The FRRs of FR-TP and P-RP have higher values than those of the SDN-ind scheme, i.e., the former two have better restoration ability than the latter one. In particular, FR-TP improves 26.89% of FRR over SDN-ind in NSFNet. This is because DA-DRSA can achieve better spectrum efficiency by using fewer spectrum resources to transmit the same services, which also contributes to the FRR. Then, with increasing of traffic load, FR-TP has better performance in FRR compared with P-RP. There are two reasons for this. (i) In general, the resources are insufficient under high load, which will hinder successful restoration, while using the SWP-LAG model to implement RSA can effectively satisfy the spectrum continuity constraint and the spectrum contiguity constraint, which can reduce spectrum fragmentation to some extent. As a result, more failure services can be accommodated again. (ii) The FR-P with triggered precomputation can guarantee the reliability of the restoration path as well as reject the unavailability of required resources on the restoration path.

Figure 11 depicts the ART under variable traffic load for NSFNet [Fig. 11(a)] and COST239 [Fig. 11(b)]. The figures

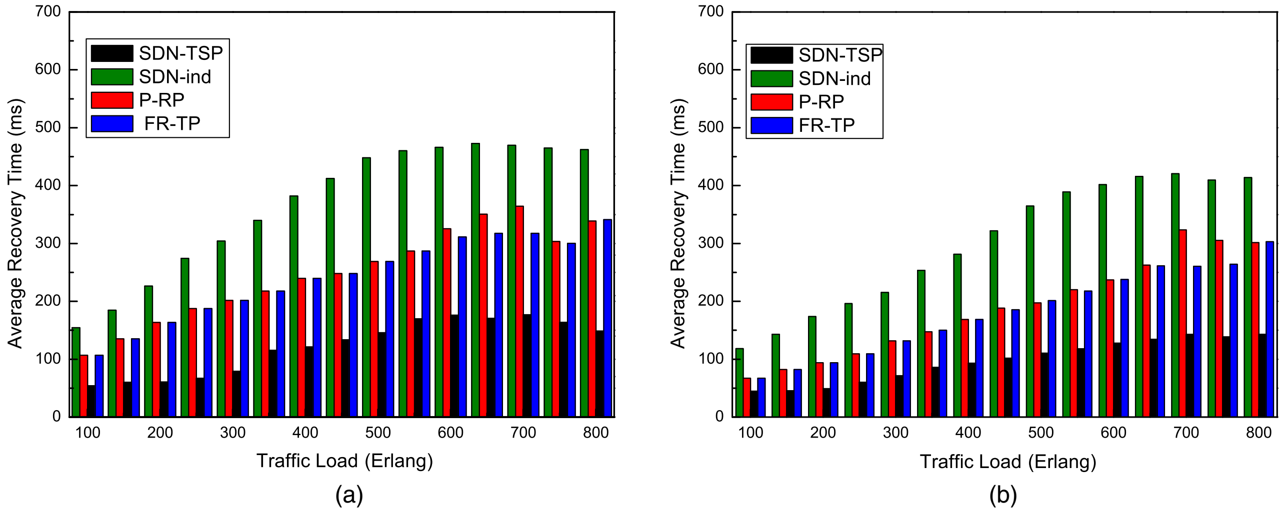


Fig. 11. ART: (a) NSFNet and (b) COST239.

show that the recovery time is strongly reduced by the proposed FR-TP strategy. The recovery time of the restoration scheme depends on the failure detection time, computation time, and provisioning time of a new restoration path. Compared with the SDN-ind scheme, since the proposed P-RP and FR-TP strategies use the precomputation restoration method to effectively reduce the complicated computation time and optical node pre-configuration, this greatly reduces the recovery time after the failure occurs. As shown in Fig. 11, the ART of FR-TP can improve by 30.4% over SDN-ind under 650 Erlangs in NSFNet, while improvement is 33.10% in COST239. The ARTs of FR-TP and P-RP have almost the same performance, but FR-TP has relatively lower restoration time than the P-RP scheme under overload. For example, under 700 Erlangs, the ART of FR-TP improves by 12.9% and 17.62% in both the NSFNet and COST239 networks, respectively. This is because P-RP cannot guarantee the availability of the required resources on the restoration path. At the moment of instantiation of the restoration path when a link fails, a new restoration path needs to be computed and then installed if the required resources are no longer available. Considering the entire process, the restoration time of P-RP will increase.

VII. CONCLUSION

In this paper, to guarantee QoS while optimizing capacity efficiency when a lightpath is interrupted, we propose a fast restoration strategy based on triggered precomputation in software-defined elastic optical inter-datacenter networks. The FR-TP strategy can not only avoid the path computation time, but also guarantees the availability of the required recovery resources. As a result, we gain better recovery time and recovery success probability. Meanwhile, we also propose a SWP-LAG model that performs dynamic distance-adaptive routing and spectrum allocation for multi-granularity services. This resource optimization method effectively achieves one-step DRSA and satisfies the two constraints of EONs with higher spectrum efficiency and lower blocking probability. Simulation results indicate that our failure restoration strategy can provide reliable and fast restoration under an extended OF-based control plane framework. Compared with SDN-TSP and SDN-ind, the proposed FR-TP strategy can restore the interrupted services quickly by using the precomputed path, with lower network resource overhead.

ACKNOWLEDGMENT

This work was made possible with funding from the National Natural Science Foundation of China (61401052), the Project of China Scholarship Council (201608500030), the Science and Technology Project of Chongqing Municipal Education Commission (KJ1400418, KJ1500445), the Starting Foundation for Doctors of Chongqing University of Posts and Telecommunications (A2015-09), and the Program for Innovation Team Building at Institutions of Higher Education in Chongqing (CXTDX201601020).

REFERENCES

[1] P. Lu, L. Zhang, X. Liu, J. Yao, and Z. Zhu, “Highly efficient data migration and backup for big data applications in elastic optical inter-data-center networks,” *IEEE Netw.*, vol. 29, no. 5, pp. 36–42, 2015.

[2] H. Yang, J. Zhang, Y. Zhao, H. Li, S. Huang, Y. Ji, J. Han, Y. Lin, and Y. Lee, “Cross stratum resilience for OpenFlow-enabled data center interconnection with flexi-grid optical networks,” *Opt. Switching Netw.*, vol. 11, pp. 72–82, 2014.

[3] “Cisco Visual Networking Index: Forecast and Methodology, 2014–2019,” Cisco White Paper, 2015.

[4] W. Zhang, H. Wang, and K. Bergman, “Next-generation optically-interconnected high-performance data centers,” *J. Lightwave Technol.*, vol. 30, no. 24, pp. 3836–3844, 2012.

[5] O. Gerstel, M. Jinno, A. Lord, and S. J. B. Yoo, “Elastic optical networking: A new dawn for the optical layer?” *IEEE Commun. Mag.*, vol. 50, no. 2, pp. s12–s20, 2012.

[6] Y. Xiong, X. Fan, and S. Liu, “Fairness enhanced dynamic routing and spectrum allocation in elastic optical networks,” *IET Commun.*, vol. 10, no. 9, pp. 1012–1020, 2016.

[7] M. Jinno, H. Takara, B. Kozicki, Y. Tsukishima, Y. Sone, and S. Matsuoka, “Spectrum-efficient and scalable elastic optical path network: Architecture, benefits, and enabling technologies,” *IEEE Commun. Mag.*, vol. 47, no. 11, pp. 66–73, 2009.

[8] H. Takara, A. Sano, T. Kobayashi, H. Kubota, H. Kawakami, A. Matsuura, Y. Miyamoto, Y. Abe, H. Ono, K. Shikama, Y. Goto, K. Tsujikawa, Y. Sasaki, I. Ishida, K. Takenaga, S. Matsuo, K. Saitoh, M. Koshiba, and T. Morioka, “1.01-Pb/s (12 SDM/222 WDM/456 Gb/s) crosstalk-managed transmission with 91.4-b/s/Hz aggregate spectral efficiency,” in *European Conf. Exhibition on Optical Communication*, Amsterdam, The Netherlands, 2012.

[9] L. Liu, W. R. Peng, R. Casellas, T. Tsuritani, and R. Muñoz, “Design and performance evaluation of an OpenFlow-based control plane for software-defined elastic optical networks with direct-detection optical OFDM (DDO-OFDM) transmission,” *Opt. Express*, vol. 22, no. 1, pp. 30–40, 2014.

[10] M. Channegowda, R. Nejabati, M. Rashidi Fard, S. Peng, N. Amaya, G. Zervas, D. Simeonidou, R. Vilalta, R. Casellas, R. Martínez, R. Muñoz, L. Liu, T. Tsuritani, I. Morita, A. Autenrieth, J. P. Elbers, P. Kostecki, and P. Kaczmarek, “Experimental demonstration of an OpenFlow based software-defined optical network employing packet, fixed and flexible DWDM grid technologies on an international multi-domain testbed,” *Opt. Express*, vol. 21, no. 5, pp. 5487–5498, 2013.

[11] N. Amaya, S. Yan, M. Channegowda, B. R. Rofoee, Y. Shu, M. Rashidi, Y. Ou, E. Hugues-Salas, G. Zervas, R. Nejabati, D. Simeonidou, B. J. Puttnam, W. Klaus, J. Sakaguchi, T. Miyazawa, Y. Awaji, H. Harai, and N. Wada, “Software defined networking (SDN) over space division multiplexing (SDM) optical networks: Features, benefits and experimental demonstration,” *Opt. Express*, vol. 22, no. 3, pp. 3638–3647, 2014.

[12] S. Das, G. Parulkar, and N. McKeown, “Why OpenFlow/SDN can succeed where GMPLS failed,” in *European Conf. Exhibition on Optical Communications*, Amsterdam, The Netherlands, 2012.

[13] F. Paolucci, F. Cugini, N. Hussain, F. Fresi, and L. Poti, “OpenFlow-based flexible optical networks with enhanced monitoring functionalities,” in *European Conf. Exhibition on Optical Communications*, Amsterdam, The Netherlands, 2012.

[14] L. Liu, R. Muñoz, R. Casellas, T. Tsuritani, and I. Morita, "OpenSlice: An OpenFlow-based control plane for spectrum sliced elastic optical path networks," *Opt. Express*, vol. 21, no. 4, pp. 4194–4204, 2013.

[15] H. Yang, J. Zhang, Y. Zhao, Y. Ji, H. Li, Y. Lin, G. Li, J. Han, Y. Lee, and T. Ma, "Performance evaluation of time-aware enhanced software defined networking (TeSDN) for elastic data center optical interconnection," *Opt. Express*, vol. 22, no. 15, pp. 17630–17643, 2014.

[16] H. Yang, J. Zhang, Y. Zhao, Y. Ji, J. Wu, X. Yu, J. Han, and Y. Lee, "Global resources integrated resilience for software defined data center interconnection based on IP over elastic optical network," *IEEE Commun. Lett.*, vol. 18, no. 10, pp. 1735–1738, 2014.

[17] L. Liu, W. Peng, R. Casellas, T. Tsuritani, and I. Morita, "Experimental demonstration of OpenFlow-based dynamic restoration in elastic optical networks on GENI testbed," in *European Conf. Optical Communication (ECOC)*, Cannes, France, 2014, pp. 1–3.

[18] L. Liu, H. Y. Choi, T. Tsuritani, I. Morita, R. Casellas, R. Martínez, and R. Muñoz, "First proof-of-concept demonstration of OpenFlow-controlled elastic optical networks employing flexible transmitter/receiver," in *Int. Conf. Photonic Switching*, Corsica, France, 2012.

[19] A. Giorgetti, F. Paolucci, F. Cugini, and P. Castoldi, "Fast restoration in SDN-based flexible optical networks," in *Optical Fiber Communication Conf.*, San Francisco, CA, 2014.

[20] L. Liu, W. R. Peng, R. Casellas, and T. Tsuritani, "Dynamic OpenFlow-based lightpath restoration in elastic optical networks on the GENI testbed," *J. Lightwave Technol.*, vol. 33, no. 8, pp. 1531–1539, 2015.

[21] A. Giorgetti, F. Paolucci, F. Cugini, and P. Castoldi, "Dynamic restoration with GMPLS and SDN control plane in elastic optical networks," *J. Opt. Commun. Netw.*, vol. 7, no. 2, pp. A174–A182, 2015.

[22] A. Giorgetti, N. Sambo, I. Cerutti, N. Andriolli, and P. Castoldi, "Label preference schemes for lightpath provisioning and restoration in distributed GMPLS networks," *J. Lightwave Technol.*, vol. 27, no. 6, pp. 688–697, 2009.

[23] Z. Zhu, W. Lu, L. Zhang, and N. Ansari, "Dynamic service provisioning in elastic optical networks with hybrid single-/multi-path routing," *J. Lightwave Technol.*, vol. 31, no. 1, pp. 15–22, 2013.

[24] S. Talebi, I. Katib, and G. N. Rouskas, "Distance-adaptive routing and spectrum assignment in rings," *IET Netw.*, vol. 5, no. 3, pp. 64–70, 2016.

[25] S. Talebi and G. N. Rouskas, "On distance-adaptive routing and spectrum assignment in mesh elastic optical networks," *J. Opt. Commun. Netw.*, vol. 9, no. 5, pp. 456–465, 2017.

[26] K. Christodoulopoulos, I. Tomkos, and E. A. Varvarigos, "Elastic bandwidth allocation in flexible OFDM-based optical networks," *J. Lightwave Technol.*, vol. 29, no. 9, pp. 1354–1366, 2011.

[27] Y. Xiong, Y. Li, X. Dong, Y. Gao, and G. N. Rouskas, "Exploiting SDN principles for extremely fast restoration in elastic optical datacenter networks," in *Global Communications Conf. (GLOBECOM)*, Washington, DC, 2016.

[28] A. Cai, G. Shen, L. Peng, and M. Zukerman, "Novel node-arc model and multiiteration heuristics for static routing and spectrum assignment in elastic optical networks," *J. Lightwave Technol.*, vol. 31, no. 21, pp. 3402–3413, 2013.

[29] G. Shen, S. K. Bose, T. H. Cheng, C. Lu, and T. Y. Chai, "Efficient heuristic algorithms for light-path routing and wavelength assignment in WDM networks under dynamically varying loads," *Comput. Commun.*, vol. 24, no. 3, pp. 364–373, 2001.

[30] A. Bocoi, M. Schuster, F. Rambach, M. Kiese, C. Bunge, and B. Spinnler, "Reach-dependent capacity in optical networks enabled by OFDM," in *Optical Fiber Communications Conf./National Fiber Optic Engineers Conf. (OFC/NFOFC)*, San Diego, CA, 2009.

[31] C. Ma, J. Zhang, Y. Zhao, Z. Jin, Y. Shi, Y. Wang, and M. Yin, "Bandwidth-adaptability protection with content connectivity against disaster in elastic optical datacenter networks," *Photon. Netw. Commun.*, vol. 30, no. 2, pp. 309–320, 2015.