# Dynamic Cross-Layer Restoration to Resolve Packet Layer Outages in FlexE-over-EONs

Meihan Wu, Nelson L. S. da Fonseca, and Zuqing Zhu, *Senior Member, IEEE*

***Abstract*—As a promising technology, Flex Ethernet (FlexE) helps to realize deterministic and ultra-low latency in metro and transport networks. Meanwhile, previous studies have confirmed the advantages of the symbiosis of FlexE and elastic optical network (EON) (*i.e.*, a FlexE-over-EON) on resource utilization and cost-effectiveness. In this paper, we consider the cross-layer restoration (CLR) in FlexE-over-EONs based on the FlexE-aware architecture. Specifically, we address the situation where an outage happened on one FlexE switch in the packet layer to bring it offline temporarily and then the affected client flows need to be recovered quickly and proactively. Three CLR strategies are first proposed to fully explore the flexibility of FlexE-over-EON for restoring the affected flows. Then, with the strategies, we formulate an integer linear programming (ILP) model and design an auxiliary graph (AG) based algorithm to reroute the affected flows as well as minimize the additional operational expense (OPEX) incurred during the CLR. Extensive simulations verify the effectiveness of our proposed CLR algorithms.**

***Index Terms*—Flexible Ethernet (FlexE), Elastic optical networks (EONs), Cross-layer restoration, Network survivability.**

## I. Introduction

NOWADAYS, the global deployment of 5G networks and cloud computing resulted in not only the rapid growth of network services and traffic but also more stringent quality-of-service (QoS) demands [1, 2]. For instance, many emerging network services, such as autonomous driving, virtual reality, and remote surgery, require deterministic and ultra-low latency/jitter as well as ultra-high reliability [3, 4]. This has stimulated intensive research and development (R&D) in multiple areas, including datacenter (DC) networking, edge/cloud computing, and time-sensitive networking. Among the R&D advances, an important one is the Flex Ethernet (FlexE) [5] released by the Optical Internetworking Forum (OIF), because it is a promising technology to achieve deterministic and ultra-low latency/jitter in metro and transport networks [6]. FlexE defines new Ethernet connection types to enable network operators to utilize the bandwidth resources in underlying optical networks adaptively, and its interface techniques can realize hard-isolation for network services [7], which is essential for QoS-guaranteed network slicing [8–10].

By using time-division multiplexing (TDM), FlexE supports various media access control (MAC) rates that are not restricted by the existing physical channel (PHY) of Ethernet [5]. Specifically, the most recent implementation agreement of FlexE (FlexE 2.1) promises to be able to carry the collections of 50 GbE, 100 GbE, 200 GbE, and 400 GbE PHYs [5]. With these PHYs, FlexE inserts a shim layer between the MAC and physical layers of Ethernet. The shim layer leverages TDM to divide the bandwidth of each group of PHYs into a series of calendar slots (CS'), and allocates the CS' to the MAC interfaces from FlexE clients to match their data-rates (as shown in the right subfigure of Fig. 1). Note that, although there are only 5 types of PHY data-rates supported by FlexE 2.1, various MAC data-rates can be mapped to PHYs.

Specifically, the data-rate mismatch can be resolved by utilizing three mechanisms, *i.e.*, bonding, sub-rating and channelization [5]. With bonding, FlexE can support one MAC interface with multiple PHYs, *e.g.*, bonding two PHYs of 50 GbE and 100 GbE, respectively, to carry a 150 Gbps MAC interface. In contrast, sub-rating helps to support the MAC interfaces whose data-rates are lower than that of a PHY, *e.g.*, transmitting the traffic from a 25 Gbps MAC interface over a 50 GbE PHY. Finally, channelization provides the flexibility to map multiple MAC interfaces to a combination of PHYs, *e.g.*, supporting two MAC interfaces of 50 Gbps and 150 Gbps, respectively, with two bonded 100 GbE PHYs.

Meanwhile, to facilitate fiber transmissions, operators need to further map the PHYs of FlexE to the transport boxes (T-Boxes) in the optical layer [11–14]. This makes the planning and provisioning of FlexE-over-Optical networks intrinsically different from and much more difficult than those of conventional Packet-over-Optical networks, because complex cross-layer mappings (*i.e.*, first from MAC interfaces to FlexE PHYs, and then from FlexE PHYs to T-Boxes) have to be tackled. Previously, by assuming that the optical layer is based on fixed-grid wavelength-division multiplexing (WDM), the authors of [11] performed a comprehensive analysis to compare the FlexE over WDM networks (FlexE-over-WDMs) built with the FlexE-unaware, FlexE-partially-aware, and FlexE-aware architectures. Their analysis indicated that as each T-Box in the FlexE-aware architecture contains a FlexE shim layer to identify the traffic from each MAC interface in PHYs, this architecture is fully compatible with FlexE and thus performs the best in terms of resource utilization and cost-effectiveness.

Later on, in [12], we showed that by architecting the optical layer with a flexible-grid elastic optical network (EON) [15–17], one can further explore the advantages of the FlexE-aware architecture. Specifically, FlexE-over-EONs can be more beneficial than FlexE-over-WDM for the following two reasons. Firstly, with bandwidth-variable transponders (BV-Ts) and bandwidth-variable wavelength-selective switches (BV-WSS'), EONs can allocate optical spectra to lightpaths in a much more

M. Wu and Z. Zhu are with the School of Information Science and Technology, University of Science and Technology of China, Hefei, Anhui 230027, P. R. China (email: zqzhu@ieee.org).

N. Fonseca is with the Institute of Computing, State University of Campinas, Campinas, SP 13083-852, Brazil.

Manuscript received on November 19, 2021.

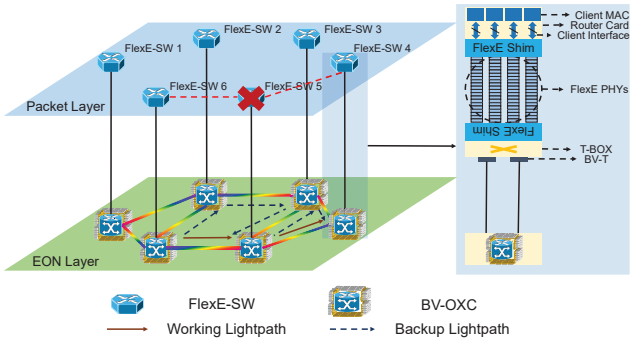



Fig. 1. Architecture of FlexE-over-EON.

flexible manner (*i.e.*, with a granularity of 12.5 GHz or even narrower) [18]. Secondly, by equipping sliceable BV-Ts in T-Boxes [19–21], an operator can redistribute the capacity of a T-Box to its BV-Ts dynamically during operation.

Even though previous studies have considered the planning and provisioning of both FlexE-over-WDMs and FlexE-over-EONs [11–14], the restoration schemes to address the network failures in them have not been fully explored yet. Note that, if the survivability of these networks is not properly maintained to secure their data transport capability consistently, the QoS demand of ultra-high reliability can never be satisfied. For the optical layer, various restoration schemes have been designed to ensure its survivability [22–26]. However, a recent analysis by Google showed that in the company's wide-area networks (WANs) (*i.e.*, B2 and B4), packet layer outages happened much more frequently than failures in the optical layer [27].

Note that, we cannot leverage the restoration schemes in the optical layer to address packet layer outages [28], which can be seen in the example in Fig. 1. Specifically, Fig. 1 shows an FlexE-over-EON, where the EON layer sets up lightpaths to carry traffic from the packet layer, and the FlexE-based packet layer routes packet flows coming in from the MAC interfaces of FlexE clients over the lightpaths. We can see that the outage on FlexE-based packet switch (FlexE-SW) 5 will disrupt all the traffic passing through it for the communications between *FlexE-SWs* 6 and 4, even though the lightpaths that connect *FlexE-SWs* 6 and 5 and *FlexE-SWs* 5 and 4 have been secured with 1+1 protection in the EON layer.

Although the affected flows will eventually be recovered when the packet layer detects the outage and updates their routing paths accordingly, the restoration latency can be too long (*e.g.*, it can take a while just for the related routing tables to converge) to violate the service-level agreements (SLAs) of the emerging network services that are sensitive to latency and reliability. Therefore, we need a proactive cross-layer restoration (CLR) scheme to properly resolve the packet layer outages in FlexE-over-EONs, which, to the best of our knowledge, has not been considered in the literature yet.

In this work, we study how to realize CLR to resolve the packet layer outages in FlexE-over-EONs proactively and cost-effectively. Specifically, we consider the case in which a single node failure happens in the packet layer during the operation of a FlexE-over-EON, and propose CLR algorithms to quickly recover the affected packet flows and minimize the operational expenses (OPEX) generated during the CLR process. The proposed algorithms are "cross-layer" ones because they jointly consider packet flow rerouting together with EON layer reconfiguration to restore the affected flows. Specifically, each affected flow can be recovered by 1) rerouting it over existing lightpath(s), 2) reconfiguring existing lightpath(s) with spectrum expansion [20] to allocate new capacity for it, and 3) setting up new lightpath(s) to carry it. The first strategy tries to leverage the spare capacity in the packet layer, while the last two strategies explore the spare spectra in the EON layer. As for the OPEX introduced by the CLR, we consider the additional usages of BV-Ts, T-Boxes, and frequency slots, and the cost of lightpath reconfiguration operations.

Based on the considerations mentioned above, we first formulate an integer linear programming (ILP) model to optimize the CLR scheme of affected flows with the three strategies, to minimize the OPEX of the CLR. As the ILP can obtain optimal solutions for small-scale CLR problems, we use it as an accountable benchmark to evaluate the performance of our other algorithms. Then, we design a time-efficient heuristic based on auxiliary graph (AG), to solve large-scale CLR problems quickly. The performance of the proposed algorithms are evaluated with extensive simulations. Our simulation results show that our proposed AG-based algorithm can approximate the optimal results from the ILP and outperform a greedy-based benchmark significantly.

The rest of the paper is as follows. Section II surveys the related work. We describe the network model of the CLR in a FlexE-over-EON in Section III. The ILP model and AG-based heuristic for optimizing the CLR are presented in Sections IV and V, respectively. In Section VI, we discuss the performance evaluations. Finally, Section VII summarizes the paper.

## II. Related Work

Ethernet is one of the most widely used Layer 2 technologies to connect network elements for transmitting IP traffic in access and metro networks. Meanwhile, putting Ethernet over optical networks can effectively improve its transmission reach, and thus has become a common practice in metro and core networks [11]. The invention of FlexE helps resolve the mismatch between traditional Ethernet interfaces and standard data-rates of optical transport network (OTN) [5], making FlexE-over-Optical networks promising. Moreover, our study in [12] found that as EON has improved flexibility in the optical layer, FlexE and EON can benefit each other mutually, to realize more adaptive and cost-effective service provisioning.

The planning and provisioning of FlexE-over-Optical networks need to address multilayer scenarios. Specifically, we need to accomplish the virtual topology design in the optical layer and the traffic grooming in the FlexE-based packet layer [11–14]. The virtual topology design sets up lightpaths by solving the routing and spectrum assignment (RSA) problem [29, 30] to layout virtual links according to the traffic matrix of the packet layer. The traffic grooming mutiplexes/demultiplexes packet flows coming in from the MAC interfaces of FlexE clients over the virtual links. Note that, the unique characteristics of FlexE will make the planning and provision-

ing schemes developed for generic Packet-over-EONs [31–34] not applicable, especially for the traffic grooming part.

As for FlexE-over-Optical networks, the studies in [35, 36] analyzed their architectural advantages, and Zhu *et al.* [37] proposed a technique to enhance the security of data transfers in them. However, these studies did not address the problems of planning and provisioning FlexE-over-Optical networks. Previously, by assuming that the optical layer is based on fixed-grid WDM, people have investigated the planning and provisioning of FlexE-over-WDMs in [11, 14]. In [11], the authors tackled the multilayer planning of FlexE-over-WDMs and compared the networks architected with the FlexE-unaware, FlexE-partially-aware, and FlexE-aware architectures. Koulougli *et al.* [14] considered a multi-domain FlexE-over-WDM and designed an efficient provisioning algorithm to solve the routing and FlexE assignment problem in it. On the other hand, we have studied the planning of FlexE-over-EONs in [12, 13], and our simulation results demonstrated the advantages of FlexE-over-EONs over FlexE-over-WDMs, in terms of adaptivity and cost-effectiveness.

Although the planning and provisioning of FlexE-over-EONs over FlexE-over-WDMs have been studied in the literature, the restoration schemes to address the network failures in them have not been fully explored yet, especially for the CLR schemes that can resolve packet layer outages. Meanwhile, as we have explained in the previous section, the restoration schemes that were developed for guaranteeing the survivability of the optical layer [22–26] cannot be leveraged to resolve packet layer outages quickly and efficiently. This motivates us to investigate how to realize CLR to resolve the packet layer outages in FlexE-over-EONs proactively and cost-effectively.

To the best of our knowledge, the only existing work that addressed the restoration schemes in FlexE-over-Optical networks is that in [38]. Specifically, the authors proposed a polynomial time algorithm to restore the traffic affected by PHY failures in a FlexE-over-WDM using the spare capacity of deployed PHYs. Nevertheless, they only tried to restore affected traffic within the packet layer (*i.e.*, did not address CLR), and did not consider FlexE-over-EONs. This, however, restricts their restoration scheme's capability on resource utilization. For instance, reconfiguring an existing lightpath with spectrum expansion will not be feasible in FlexE-over-WDMs, because of the absence of sliceable BV-Ts.

## III. Cross-layer Restoration (CLR) in FlexE-over-EONs

In this section, we first describe the network model of FlexE-over-EON, and then explain the fundamentals of the three CLR strategies considered in this work.

### A. Network Model

FlexE-over-EONs can be realized with the FlexE-unaware, FlexE-partially-aware, and FlexE-aware architectures [11, 12]. Among them, the FlexE-aware architecture is the most flexible and beneficial one [12]. Therefore, similar to the work in [13], we also assume that the FlexE-over-EON under consideration is built with the FlexE-aware architecture. As shown in Fig. 2, the FlexE-aware architecture deploys FlexE shims not only between client MAC interfaces and PHYs but also between PHYs and T-Boxes, to recognize the data from each client MAC interface in the PHYs and realize flow-level routing [5].

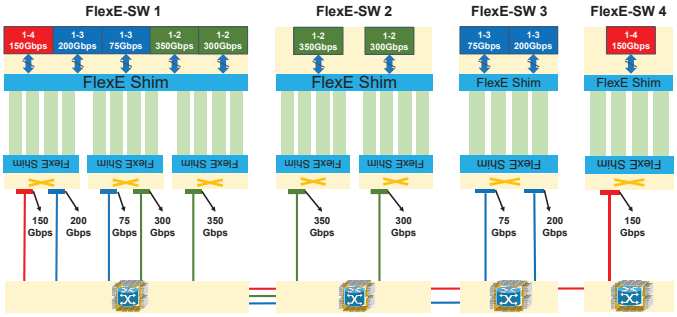



Fig. 2. FlexE-over-EON based on FlexE-aware architecture.

The network architecture of the FlexE-over-EON considered in this work has been shown in Fig. 1. We model the topology of the EON layer (*i.e.*, the physical topology of the FlexE-over-EON) as $G_o(V_o, E_o)$, where $V_o$ denotes the set of optical nodes, each of which consists of a bandwidth-variable optical cross-connect (BV-OXC), and $E_o$ represents the set of bi-directional fiber links. Meanwhile, the topology of the FlexE-based packet layer can be modeled as $G_f(V_f, E_f)$, where each node in set $V_f$ is a FlexE-based packet switch (FlexE-SW) that locally connects to a BV-OXC in $V_o$, and $E_f$ represents the set of logical links among the FlexE-SWs. Note that, each logical link $e_f \in E_f$ is actually supported by a lightpath in the EON layer. Hence, if a lightpath has been set up to connect two BV-OXCs, their local FlexE-SWs are connected by a logic link in the packet layer. In other words, if two FlexE-SWs are not directly connected in $G_f(V_f, E_f)$, there is no lightpath between the corresponding BV-OXCs in the EON layer.

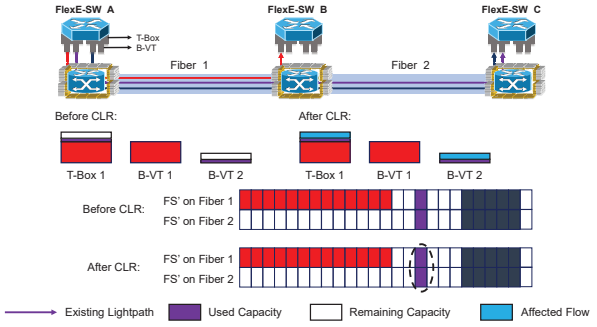
In the EON layer, each fiber link $e_o \in E_o$ can accommodate $F$ frequency slots (FS'), each of which has a bandwidth of 12.5 GHz [12]. Each FlexE-SW consists of router cards (on the client side) and T-Boxes (on the network side), where both the numbers of router cards and T-Boxes are fixed [13]. As shown in Fig. 2, each router card contains a FlexE shim to groom the traffic from the MAC interfaces of FlexE clients on PHYs, while the FlexE shim on a T-Box can recognize the traffic from each client MAC interface in the PHYs and send it to a proper BV-T in the T-Box. Here, each T-Box consists of a fixed number of BV-Ts whose total capacity is also fixed. We assume that each BV-T is a sliceable one [19–21], *i.e.*, the operator can redistribute the capacity of a T-Box to its BV-Ts dynamically with spectrum expansion.

Each BV-T uses the FS' on fiber links to establish its lightpath. For simplicity, this work assumes that all the lightpaths in the EON layer are setup with the same RSA algorithm (*e.g.*, the shortest-path routing and first-fit spectrum assignment (SP-FF) [29]). Meanwhile, when the routing path of a BV-T's lightpath in $G_o(V_o, E_o)$ has been determined, it selects the modulation format according to the length of the path [39, 40]. Specifically, we consider four modulation formats, *i.e.*, BPSK, QPSK, 8QAM, and 16QAM, whose parameters are listed in

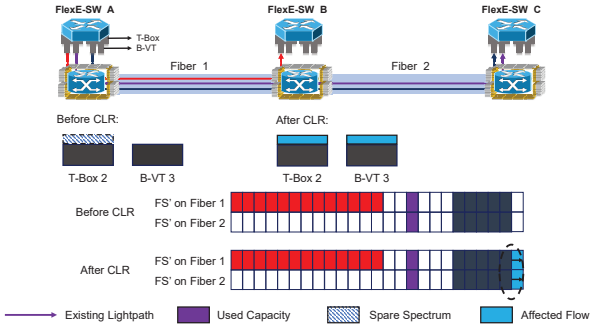
Table I (according to the experimental results in [39]).

TABLE I
PARAMETERS OF MODULATION FORMATS OF LIGHTPATHS

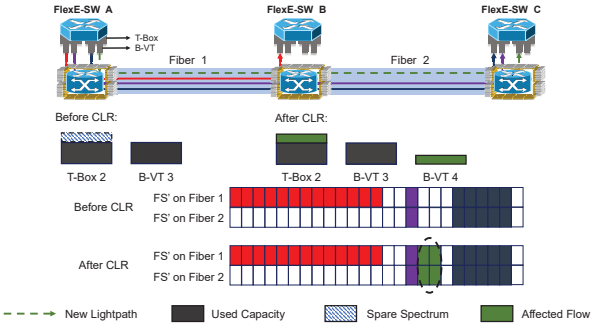
| Modulation Format | BPSK | QPSK | 8QAM | 16QAM |
|---|---|---|---|---|
| Capacity per FS (Gb/s) | 12.5 | 25 | 37.5 | 50 |
| Transmission Reach (km) | 4,800 | 2,400 | 1,200 | 600 |
| Power Usage per FS (W) | 112.4 | 133.4 | 154.4 | 175.5 |




(a) Rerouting over existing lightpaths




(b) Reconfiguring existing lightpaths with spectrum expansion




(c) Setting up new lightpaths

Fig. 3. Cross-layer restoration strategies.

Hence, each lightpath (*i.e.*, a logic link $e_f \in E_f$) can be described with a few attributes, including the source and destination FlexE-SWs, transmission capacity, modulation format, routing path, and spectrum assignment. Then, the network model of the FlexE-over-EON can be simplified as $G(V, E)$, where $V$ and $E$ denote the sets of FlexE-SWs and logical links, respectively, and the information about the EON layer is represented by the attributes of the logic links. Meanwhile, for the traffic from each client MAC interface, it is first mapped to a BV-T in the source FlexE-SW, and then routed to the destination FlexE-SW over the logic links in $G(V, E)$.

### B. CLR Strategies

For the CLR in a FlexE-over-EON, we consider the failure of a FlexE-SW $v \in V$ during network operation. After the failure, we first update the status of the FlexE-over-EON to temporarily remove the broken FlexE-SW $v$, obtain all the client traffic that uses it as an intermediate node, and store the traffic in a matrix $\mathbf{R}$. Then, $\mathbf{R}$ denotes the traffic that has been impacted by the failure and should be recovered with CLR. Meanwhile, we hope to point out that as the traffic whose source or destination node is $v$ cannot be restored until the broken FlexE-SW $v$ has been fixed, it is unrecoverable and thus should be excluded from the CLR of this work. The CLR leverages the spare resources in the packet and EON layers to quickly restore the affected traffic in $\mathbf{R}$, while the price (*i.e.*, the OPEX of the CLR) is that there will be additional usages of BV-Ts, T-Boxes, and FS' and incremental operational cost due to lightpath reconfigurations[1]. Therefore, the optimization of the CLR should try to minimize its OPEX.

Specifically, the CLR can utilize the three strategies in Fig. 3 to recover the traffic in $\mathbf{R}$. The first strategy in Fig. 3(a) restores the affected traffic by rerouting it over existing lightpaths to leverage the spare capacities there. Here, as the *BV-T* 2 in *T-Box* 1 has sufficient spare capacity to accommodate the affected flow, we can restore the affected flow by rerouting it through the existing lightpath from *BV-T* 2. Since it does not involve any operation in the EON layer, the first strategy is the fastest one, and in the meantime, it does not result in any additional OPEX because it will not occupy any new BV-Ts, T-Boxes, or FS' or reconfigure any lightpaths.

As shown in Fig. 3(b), the second strategy also reroutes the affected traffic over existing lightpaths. However, this time, the *BV-T* 3 in *T-Box* 2 does not have enough spare capacity but *T-Box* 2 still has some room for a spectral expansion. Hence, we first reconfigure *BV-T* 3 to expand the spectrum assignment of its lightpath and then reroute the affected flow over the spectrally-expanded lightpath. Note that, the spectrum expansion should be conducted under 1) the capacity constraint of each T-Box [13], and 2) the spectrum continuity, contiguity, and non-overlapping constraints [29]. This strategy does not use any new BV-Ts or T-Boxes, but it occupies new FS' after the spectrum expansion and also involves lightpath reconfigurations. As shown in Fig. 3(c), the last strategy sets up a new lightpath with the *BV-T* 4 in *T-Box* 2 to restore the affected flow. Hence, it might use new BV-Ts, T-Boxes, and FS' and will reconfigure lightpaths.

Considering the fact that over-provisioning is a common practice in today's metro and core networks, we assume that the spare resources in the FlexE-over-EON are sufficient for the CLR to recover all the affected traffic in $\mathbf{R}$ with the three

[1]Note that, as $\mathbf{R}$ only contains the transit traffic of a broken FlexE-SW, its total traffic volume might not be very large. Hence, it would be reasonable to assume that all the affected traffic in $\mathbf{R}$ can be restored with the spare resources in the FlexE-over-EON. Otherwise, if the spare resources are indeed insufficient, we can adopt a pre-processing method to select the recoverable flows accordingly. Then, $\mathbf{R}$ can be understood as the traffic matrix of the selected recoverable flows. Furthermore, in our future work, we will formulate a new optimization that jointly considers the selection of recoverable flows and the planning of their restoration, for more thoroughly addressing the situation where not all the affected flows in $\mathbf{R}$ can be recovered due to limited resources.

aforementioned strategies. Then, the OPEX of the CLR can be quantified as follows. First of all, we quantify the cost of the additional usages of BV-Ts, T-Boxes, and FS' based on their power consumptions. Specifically, we denote the static power consumptions of a new BV-T and a new T-Box as $w_b$ and $w_t$, respectively, and the power consumption per new FS is represented as $w_m$, where we use $m = \{1, 2, 3, 4\}$ to respectively indicate that the modulation format of the FS is BPSK, QPSK, 8QAM, or 16QAM. Meanwhile, we assume that the cost of one lightpath reconfiguration is $c_r$. Therefore, we can obtain the unit OPEX of the three CLR strategies as

$$c = \begin{cases} 0, & \text{First Strategy,} \\ c_w \cdot w_m \cdot n_f + c_r, & \text{Second Strategy,} \\ c_w \cdot (w_m \cdot n_f + 2 \cdot w_b + n_t \cdot w_t) + c_r, & \text{Third Strategy,} \end{cases} \tag{1}$$

where $c_w$ is the unit cost of power consumption, $n_f$ is the number of newly-used FS', and $n_t \in [0, 2]$ is the number of new T-Boxes that will be needed to set up a new lightpath. Eq. (1) indicates that the third strategy is the most expensive, followed by the second one, while the first one is the cheapest because it does not use any additional BV-Ts, T-Boxes, or FS'.

## IV. Integer Linear Programming Model

In this section, we formulate an ILP model to optimize the CLR in a FlexE-over-EON with the three strategies discussed in the previous section, such that all the traffic that is affected by the outage of a FlexE-SW can be recovered successfully and the OPEX of the CLR is minimized.

**Notations:**

- $G(V, E)$: the topology of the FlexE-over-EON, where $V$ denotes the set of FlexE-SWs and $E$ represents the set of logic links in the packet layer.
- $e$: a logical link that represents a lightpath and its attributes of capacity, modulation format, routing path, and spectrum assignment. We use $e = (u, v) \in E$ to refer to a logic link between FlexE-SWs $u$ and $v$.
- $\mathbf{R}$: the matrix of affected traffic. Here, each element in $\mathbf{R}$ denotes an affected client flow that should be recovered in the CLR, and it can be referred to as $r_i = (s_i, d_i, x_i)$, where $i$ is the unique index, $s_i$ and $d_i$ are its source and destination, respectively, and $x_i$ represents its bit-rate.
- $g_{u,v}$: the indicator that equals 1 if after the outage of a FlexE-SW $\tilde{v}$, lightpaths still exist or can be set up (*i.e.*, feasible) between FlexE-SWs $u$ and $v$, and 0 otherwise.
- $B_u$: the set of BV-Ts on FlexE-SW $u$.
- $T_u$: the set of T-Boxes on FlexE-SW $u$.
- $B_{u,t}$: the set of BV-Ts in a T-Box $t \in T_u$.
- $C_t$: the remaining capacity of a T-Box $t$.
- $F$: the number of FS' on each fiber link in the EON layer.
- $p_{u,v}$: the capacity per FS on each lightpath between FlexE-SWs $u$ and $v$. For the lightpath, $p_{u,v}$ is determined by its modulation format, which is selected according to its transmission distance (as listed in Table I).
- $W_{u,v}$: the power usage per FS on each lightpath between FlexE-SWs $u$ and $v$. We determine it similarly as $p_{u,v}$.
- $w'_{u,b,v}$: the start index of the FS' used by a BV-T $b \in B_u$ for a lightpath to FlexE-SW $v$, before the CLR.
- $z'_{u,b,v}$: the end index of the FS' used by a BV-T $b \in B_u$ for a lightpath to FlexE-SW $v$, before the CLR.
- $h'_{u,b}$: the indicator that equals 1 if a BV-T $b \in B_u$ is used before the CLR, and 0 otherwise.
- $h'_{u,t}$: the indicator that equals 1 if a T-Box $t \in T_u$ is used before the CLR, and 0 otherwise.
- $h'_{u,b,v}$: the indicator that equals 1 if a BV-T $b \in B_u$ is used to support a lightpath to FlexE-SW $v$ before the CLR, and 0 otherwise.
- $C_{u,b}$: the remaining capacity of a BV-T $b \in B_u$.
- $k_{e,e'}$: the indicator that equals 1 if the lightpaths of logical links $e$ and $e'$ share fiber link(s), and 0 otherwise.
- $w_b$: the static power consumption of each BV-T.
- $w_t$: the static power consumption of each T-Box.
- $c_w$: the unit cost of power consumption.
- $c_r$: the cost of one lightpath reconfiguration.

**Variables:**

- $\alpha^i_{u,b}$: the boolean variable that equals 1 if the affected flow $r_i \in \mathbf{R}$ is transmitted over a BV-T $b \in B_u$ after the CLR, and 0 otherwise.
- $\alpha^i_{u,v}$: the boolean variable that equals 1 if the flow $r_i \in \mathbf{R}$ uses a logic link between FlexE-SWs $u$ and $v$ after the CLR, and 0 otherwise.
- $\beta_{u,b,v}$: the boolean variable that equals 1 if a BV-T $b \in B_u$ for a lightpath to FlexE-SW $v$ is selected for spectrum expansion in the CLR, and 0 otherwise.
- $\delta^w_{u,b,v}$: the number of new FS' that the spectrum expansion on a BV-T $b \in B_u$ (for a lightpath to FlexE-SW $v$) extends to the lower end of the spectrum usage.
- $\delta^z_{u,b,v}$: the number of new FS' that the spectrum expansion on a BV-T $b \in B_u$ (for a lightpath to FlexE-SW $v$) extends to the upper end of the spectrum usage.
- $w_{u,b,v}$: the start index of the FS' used by a BV-T $b \in B_u$ for a lightpath to FlexE-SW $v$, after the CLR.
- $z_{u,b,v}$: the end index of the FS' used by a BV-T $b \in B_u$ for a lightpath to FlexE-SW $v$, after the CLR.
- $\gamma_{u,b,v}$: the boolean variable that equals 1 if a BV-T $b \in B_u$ is used to set up a new lightpath to FlexE-SW $v$ after the CLR, and 0 otherwise.
- $Y_{u,b,v}$: the number of FS' used by the BV-T $b \in B_u$ for a new lightpath to FlexE-SW $v$, after the CLR.
- $h_{u,b}$: the boolean variable that equals 1 if a BV-T $b \in B_u$ is used after the CLR, and 0 otherwise.
- $h_{u,b,v}$: the boolean variable that equals 1 if a BV-T $b \in B_u$ is used to support a lightpath to FlexE-SW $v$ after the CLR, and 0 otherwise.
- $h_{u,t}$: the boolean variable that equals 1 if a T-Box $t \in T_u$ is used after the CLR, and 0 otherwise.
- $o_{b,b'}$: the boolean variable that equals 1 if for two BV-Ts $b$ and $b'$ respectively of lightpaths $(u, v)$ and $(u', v')$, we have $w_{u,b,v} \leq w_{u',b',v'}$ after the CLR, and 0 otherwise.

**Objective:**

The objective is to minimize the OPEX of the CLR, as

$$\text{Minimize} \quad C, \tag{2}$$

where the OPEX $C$ can be calculated as

$$\begin{aligned} C =& c_r \cdot \sum_{u\in V}\sum_{b\in B_u}\sum_{v\in V}(\beta_{u,b,v}+\gamma_{u,b,v}) \\ &+c_w \cdot \sum_{u\in V}\sum_{b\in B_u}\sum_{v\in V}\left(\delta^w_{u,b,v}+\delta^z_{u,b,v}+Y_{u,b,v}\right)\cdot W_{u,v} \\ &+c_w \cdot \sum_{u\in V}\sum_{t\in T_u}\left(h_{u,t}-h'_{u,t}\right)\cdot w_t \cdot 2 \\ &+c_w \cdot \sum_{u\in V}\sum_{b\in B_u}\sum_{v\in V}\gamma_{u,b,v}\cdot w_b \cdot 2. \end{aligned} \tag{3}$$

On the right side of Eq. (3), the first term is for the total cost generated by lightpath reconfigurations, the second one denotes the total cost of the power consumption of the new FS' allocated in the CLR, and the last two terms represent the total power consumption cost of the new BV-Ts and T-Boxes used in the CLR. As lightpath reconfigurations can increase the CLR latency significantly, we set $c_r \gg c_w$ to ensure that the ILP tries to first minimize lightpath reconfigurations and then reduce the cost due to additional power consumption.

**Constraints:**

$$\sum_{v\in V}\alpha^i_{u,v}-\sum_{v\in V}\alpha^i_{v,u}=\begin{cases}1, & s_i=u,\\ -1, & d_i=u,\\ 0, & \text{others},\end{cases}\quad \{i:r_i\in\mathbf{R}\}, \tag{4}$$

$$\sum_{v\in V}\alpha^i_{u,v}\le 1,\quad \{i:r_i\in\mathbf{R}\},\ \forall u\in V, \tag{5}$$

$$\sum_{u\in V}\alpha^i_{u,v}\le 1,\quad \{i:r_i\in\mathbf{R}\},\ \forall v\in V. \tag{6}$$

Eqs. (4)-(6) ensure that we use single-path routing to restore each affected flow $r_i$, and along its new routing path, the flow is always transmitted over a single lightpath.

$$\alpha^i_{u,v}\le g_{u,v},\quad \{i:r_i\in\mathbf{R}\},\ \{u,v:u\neq v,\ u,v\in V\}. \tag{7}$$

Eq. (7) ensures that if the CLR of an affected flow $r_i$ uses a logic link $(u,v)$, the corresponding lightpath is feasible.

$$\sum_{b\in B_u}\alpha^i_{u,b}=\sum_{v\in V}\alpha^i_{u,v},\quad \{i:r_i\in\mathbf{R}\},\ \forall u\in V. \tag{8}$$

Eq. (8) ensures that if the CLR of an affected flow $r_i$ uses a FlexE-SW $u$, the flow has to be sent out by a BV-T $b\in B_u$.

$$\alpha^i_{u,b}\le h_{u,b},\quad \{i:r_i\in\mathbf{R}\},\ \forall u\in V,\ b\in B_u. \tag{9}$$

Eq. (9) ensures that if the CLR of an affected flow $r_i$ uses a BV-T $b\in B_u$, the BV-T is marked as used.

$$\alpha^i_{u,b}\le\frac{\sum_{v\in V}\alpha^i_{u,v}+h_{u,b,v}}{2},\ \{i:r_i\in\mathbf{R}\},\ \forall u\in V,\ b\in B_u. \tag{10}$$

Eq. (10) ensures that all the affected flows using a BV-T $b\in B_u$, whose lightpath is to FlexE-SW $v$, are transmitted to $v$.

$$\begin{cases}\beta_{u,b,v}\le\delta^w_{u,b,v}+\delta^z_{u,b,v}\le F\cdot\beta_{u,b,v},\\ w_{u,b,v}=w'_{u,b,v}-\delta^w_{u,b,v}\ge 1,\\ z_{u,b,v}=z'_{u,b,v}+\delta^z_{u,b,v}\ge 1,\end{cases} \forall v\in V,\ \{u,b:u\in V,\ b\in B_u,\ h'_{u,b}=1\}. \tag{11}$$

Eq. (11) ensures that the related variables are updated correctly, when CLR selects a BV-T $b\in B_u$ for spectrum expansion.

$$\begin{cases}\gamma_{u,b,v}\le Y_{u,b,v}\le F\cdot\gamma_{u,b,v},\\ \gamma_{u,b,v}\le w_{u,b,v}\le F\cdot\gamma_{u,b,v},\\ z_{u,b,v}=w_{u,b,v}+Y_{u,b,v}-\gamma_{u,b,v},\end{cases} \forall v\in V,\ \{u,b:u\in V,\ b\in B_u,\ h'_{u,b}=0\}. \tag{12}$$

Eq. (12) ensures that the related variables are updated correctly, when CLR sets up a new lightpath with a BV-T $b\in B_u$.

$$\gamma_{u,b,v}\le g_{u,v},\quad \forall u,v\in V,\ b\in B_u. \tag{13}$$

Eq. (13) ensures that a new lightpath to $v$ with a BV-T $b\in B_u$ can only be set up when a lightpath for $u$-$v$ is feasible.

$$\sum_{v\in V}\gamma_{u,b,v}\le 1,\quad \forall u\in V,\ b\in B_u. \tag{14}$$

Eq. (14) ensures that a new lightpath established with a BV-T $b\in B_u$ has only one destination.

$$\begin{cases}h_{u,b}=\sum_{v\in V}\gamma_{u,b,v}+h'_{u,b},\\ h_{u,b,v}=\gamma_{u,b,v}+h'_{u,b,v},\\ h_{u,t}\le h'_{u,t}+\left(1-h'_{u,t}\right)\cdot\sum_{b\in B_{u,t}}h_{u,b},\ \forall u,v\in V,\ b\in B_u.\\ h_{u,t}\ge h'_{u,t}+\frac{1-h'_{u,t}}{2}\cdot\sum_{b\in B_{u,t}}h_{u,b}\end{cases} \tag{15}$$

Eq. (15) ensures that the related variables are updated correctly after the CLR.

$$\begin{aligned}\sum_{\{i:r_i\in\mathbf{R}\}}\alpha^i_{u,b}\cdot x_i+C_{u,b}\le&\sum_{v\in V}(z_{u,b,v}-w_{u,b,v})\cdot p_{u,v}\\ &+\sum_{v\in V}(h_{u,b,v}+\gamma_{u,b,v})\cdot p_{u,v},\\ &\forall u,v\in V,\ b\in B_u.\end{aligned} \tag{16}$$

Eq. (16) ensures that the total bit-rate of the affected flows using a BV-T $b\in B_u$ does not exceed its remaining capacity.

$$\sum_{b\in B_{u,t}}(z_{u,b,v}-w_{u,b,v}+h_{u,b,v}+\gamma_{u,b,v})\cdot p_{u,v}\le C_t,\ \forall u,v\in V. \tag{17}$$

Eq. (17) ensures that the total bit-rate of the affected flows using a T-Box $t\in T_u$ does not exceed its remaining capacity.

$$\begin{cases}\delta^w_{u,b,v},\ \delta^z_{u,b,v},\ Y_{u,b,v}\in[1,F],\\ w_{u,b,v},\ z_{u,b,v}\in[1,F],\end{cases}\quad \forall u,v\in V,\ b\in B_u. \tag{18}$$

Eq. (18) limit the ranges of the variables. In order to make the formulations in the following constraint concise and clear, we abbreviate $w_{u,b,v}$, $z_{u,b,v}$, and $h_{u,b,v}$ to $w_{e,b}$, $z_{e,b}$, and $h_{e,b}$, respectively, after defining $e=(u,v)$.

$$\begin{cases}h_{e',b'}-F\cdot o_{b,b'}\le w_{e,b}-w_{e',b'}\le F\cdot(1-o_{b,b'})-h_{e',b'},\\ z_{e',b'}-w_{e,b}\le F\cdot o_{b,b'}-h_{e',b'},\\ z_{e,b}-w_{e',b'}\le F\cdot(1-o_{b,b'})-h_{e',b'},\end{cases} \{e=(u,v),e'=(u',v'):e\neq e',\ b\neq b',\ k_{e,e'}=1, \forall u,v,u',v'\in V\}. \tag{19}$$

Eq. (19) ensures that after the CLR, all the lightpaths in the EON layer still satisfy the spectrum continuity, contiguity, and non-overlapping constraints [29].

## V. Heuristic Algorithm Design

Although the ILP model can obtain the optimal solution of CLR, its complexity can make the problem-solving time-consuming or even intractable, especially for large-scale FlexE-over-EONs. This prevents it from recovering the affected flows quickly. Therefore, we design a heuristic in this section to find a cost-effective CLR scheme time-efficiently. Intuitively, the CLR problem can be solved in a greedy way, which first sorts the affected flows in **R** according to one or more metrics and then restores them sequentially in the sorted order. Meanwhile, finding the actual restoration scheme of an affected flow involves multilayer service provisioning, which usually can be solved by building an auxiliary graph (AG) and performing routing in it [28]. However, this greedy-based approach cannot optimize the cost-effectiveness of CLR properly. Specifically, as the affected flows are handled separately, it may expand the spectrum of a lightpath repeatedly and/or set up multiple new lightpaths between a same pair of FlexE-SWs. Hence, unnecessary OPEX due to redundant lightpath reconfigurations and BV-T/T-Box usages can be caused.

This motivates us to design the AG-based heuristic to consider the affected flows in **R** jointly for avoiding unnecessary spectrum expansions and new lightpaths, *i.e.*, it will only expand the spectrum of an existing lightpath once and minimize the numbers of BV-Ts and T-Boxes for new lightpaths. As there can be multiple existing lightpaths between a pair of FlexE-SWs $u$-$v$ in the EON layer, we record their logic links in a set $E_{u,v}$. Then, each logic link $e \in E_{u,v}$ has two important attributes related to the CLR: 1) its remaining capacity (RC) $e.c$, and 2) its spare capacity upper-limit $e.\hat{c}$. Here, we define the **spare capacity upper-limit (SC-UL)** of a logic link as the RC on it after its lightpath has been spectrally expanded to the maximum extent. Next, we use $(u, v, \hat{b})$ to denote the logic link that has the largest RC in $E_{u,v}$, where $\hat{b} \in B_u$ is the BV-T of its lightpath on FlexE-SW $u$, and leverage $(u, v, \tilde{b})$ to represent the logic link that has the largest SC-UL in $E_{u,v}$, where $\tilde{b} \in B_u$ is the BV-T of its lightpath on FlexE-SW $u$.

Based on the discussions above, we first design *Algorithm* 1 to build an AG for each affected flow in **R**, and then propose *Algorithm* 2 to jointly consider the AGs for realizing cost-effective CLR. The procedure in *Algorithm* 1 explains how to build the AG for an affected flow $r_i \in \mathbf{R}$ based on the current status of the FlexE-over-EON. *Line* 1 lets the node set of the AG inherit all the nodes in $V$. Then, the for-loop checks each pair of FlexE-SWs $u$ and $v$, which can be directly connected with a lightpath in $G(V, E)$ ($g_{u,v} = 1$), and inserts a weighted link in $G_a(V_a, E_a)$ to connect them (*Lines* 2-23). Here, *Line* 3 first checks all the existing lightpath(s) between FlexE-SWs $u$ and $v$ to find those that have the largest RC and the largest SC-UL, and *Line* 4 then calculate the weighting parameter $\zeta$ based on $\{g_{u,v}\}$ with the following equation

$$\zeta = \frac{1}{1 + \sum\limits_{u,v \in V} g_{u,v}}. \tag{20}$$

Specifically, $\zeta$ decreases with the potential connectivity of $G(V, E)$ (*i.e.*, $\sum\limits_{u,v \in V} g_{u,v}$), and in other words, $\zeta$ becomes smaller when more pairs of FlexE-SWs are or can be connected

**Algorithm 1:** Constructing AG for an Affected Flow

**Input**: affected flow $r_i = (s_i, d_i, x_i) \in \mathbf{R}$, $\{f_e, e \in E\}$, and status of FlexE-over-EON $G(V, E)$.
**Output**: AG $G_a(V_a, E_a)$ with weighted links.

1 $V_a = V$;
2 **for** *each pair of FlexE-SWs* $u$ *and* $v$ *with* $g_{u,v} = 1$ **do**
3 check all the logic links in $E_{u,v}$ to find those that have the largest RC (*i.e.*, $(u, v, \hat{b})$) and the largest SC-UL (*i.e.*, $(u, v, \tilde{b})$), respectively;
4 calculate the weighting parameter $\zeta$ with Eq. (20);
5 connect $u$ and $v$ directly in $G_a(V_a, E_a)$;
6 **if** $(u, v, \hat{b}).c \geq x_i$ **then**
7 assign weight of $(u, v) \in E_a$ as $\varpi_{u,v} = \zeta^2$;
8 **else if** $(u, v, \tilde{b}).\hat{c} \geq x_i$ **then**
9 $flag = 0$;
10 **for** *each logic link* $e \in E_{u,v}$ **do**
11 **if** $e.\hat{c} \geq x_i$ *AND* $f_e = 1$ **then**
12 assign weight of $(u, v) \in E_a$ as $\varpi_{u,v} = \zeta$;
13 $flag = 1$;
14 **break**;
15 **end**
16 **end**
17 **if** $flag = 0$ **then**
18 assign weight of $(u, v) \in E_a$ as $\varpi_{u,v} = 1$;
19 **end**
20 **else**
21 assign weight of $(u, v) \in E_a$ as $\varpi_{u,v} = 1$;
22 **end**
23 **end**
24 **return** $G_a(V_a, E_a)$;

directly with lightpaths in the EON layer. Next, *Line* 5 adds a link $(u, v)$ in $G_a(V_a, E_a)$ to directly connect $u$ and $v$ there, and *Lines* 6-22 assign a weight to the link.

If we have $(u, v, \hat{b}).c \geq x_i$, the affected flow $r_i$ can be restored with the RC on an existing lightpath between $u$ and $v$ and *Line* 7 assigns the link weight as $\zeta^2$. Otherwise, if we have $(u, v, \tilde{b}).\hat{c} \geq x_i$, $r_i$ can be recovered by expanding the spectrum of an existing lightpath for $(u, v)$. Here, to minimize lightpath reconfigurations, we introduce a flag $f_e$ for each logic link $e \in E$. Specifically, we have $f_e = 1$ if the logic link has already been selected for lightpath reconfigurations to restore other affected flow(s), and 0 otherwise. Then, we check all the logic links in $E_{u,v}$ to see whether a logic link $e$ that satisfies $e.\hat{c} \geq x_i$ and $f_e = 1$ can be found (*Lines* 9-16). If yes, *Line* 12 assigns the link weight of $(u, v) \in E_a$ as $\zeta$ because the CLR can restore $r_i$ with a logic link between $u$ and $v$ without introducing an additional lightpath reconfiguration. Otherwise, *Line* 18 assigns the link weight of $(u, v) \in E_a$ as 1.

Meanwhile, if none of the logic links in $E_{u,v}$ can be used to restore $r_i$ even after spectrum expansion (*i.e.*, a new lightpath needs to be set up for $(u, v)$), *Line* 21 also assigns the link weight of $(u, v) \in E_a$ as 1. The rationale behind *Lines* 6-22 is that for a $G(V, E)$ with a better potential connectivity, we should try to avoid lightpath reconfigurations more, and *vice versa*. Therefore, the weight assignment can promote the CLR

to leverage the three strategies in Fig. 3 in sequence to achieve fast and cost-effective restoration of the affected flows in **R**.

**Algorithm 2:** AG-based Algorithm for CLR Design

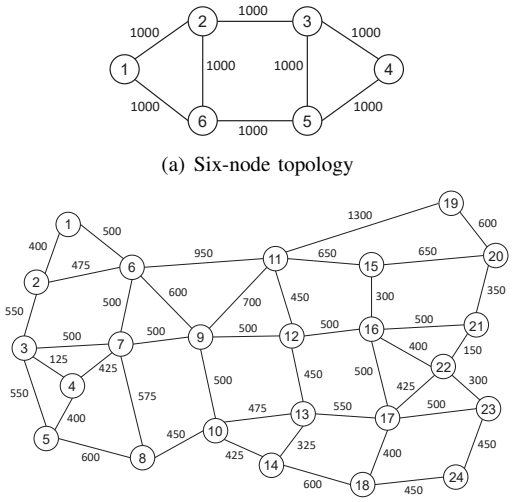
**Input**: affected flows in **R**, and status of FlexE-over-EON $G(V, E)$ after a FlexE-SW failure.
**Output**: a CLR scheme to recover all the flows in **R**.

1 sort flows in **R** in descending order of their bit-rates;
2 **for** *each flow* $r_i \in \boldsymbol{R}$ *in the sorted order* **do**
3  build the AG $G_a(V_a, E_a)$ for $r_i$ with *Algorithm* 1;
4  calculate $K$ least-weighted paths from $s_i$ to $d_i$ in $G_a(V_a, E_a)$ and store them in set $\mathcal{P}$;
5  **for** *each path* $p \in \mathcal{P}$ **do**
6  $B_p = \emptyset$;
7  **for** *each link* $(u, v)$ *on* $p$ *in sequence* **do**
8  **if** $\varpi_{u,v} = \zeta^2$ **then**
9  select a BV-T $b \in B_u$ whose logic link $e$ satisfies $e.c \geq x_i$ with first-fit;
10  insert $b$ in set $B_p$;
11  **else if** $\varpi_{u,v} = \zeta$ **then**
12  select a BV-T $b \in B_u$ whose logic link $e$ satisfies $e.\hat{c} \geq x_i$ and $f_e = 1$ with first-fit;
13  insert $\tilde{b}$ in set $B_p$;
14  **else**
15  **if** *a logic link* $e \in E_{u,v}$ *that satisfies* $e.\hat{c} \geq x_i$ *can be found* **then**
16  get the BV-T $b \in B_u$ of logic link $e$;
17  set $f_e = 1$ and insert $b$ in set $B_p$;
18  **else**
19  select an unused BV-T $b \in B_u$ and mark it for a new lightpath $e$ to $v$;
20  $f_e = 1$;
21  insert $b$ in set $B_p$ and add $e$ into $E$;
22  **end**
23  **end**
24  **end**
25  **end**
26  find the path $p^* \in \mathcal{P}$ with the smallest cost;
27  record that $r_i$ will be recovered with the BV-Ts in $B_{p^*}$ and update network status of $G(V, E)$;
28 **end**
29 aggregate the CLR schemes of all the affected flows in **R** to avoid redundant lightpath reconfigurations;
30 implement aggregated CLR schemes in FlexE-over-EON;

*Algorithm* 2 shows the overall procedure of our AG-based CLR algorithm. *Line* 1 is the initialization to sort the affected flows in **R**. Then, the for-loop of *Lines* 2-28 finds the cheapest CLR scheme of each flow in sorted order. Specifically, the CLR scheme of an affected flow $r_i$ is obtained in three steps:

1) building an AG $G_a(V_a, E_a)$ for $r_i$ with *Algorithm* 1 and calculating $K$ least-weighted paths in the AG (*Lines* 3-4),
2) checking each of the obtained paths in set $\mathcal{P}$ to get the CLR scheme of $r_i$ with it (*Lines* 5-25),
3) finding the path $p^* \in \mathcal{P}$, with which $r_i$ can be restored with the smallest cost (*Lines* 26-27).

In Step 2, we use set $B_p$ to store the BV-Ts that should be utilized to restore $r_i$ on a path $p \in \mathcal{P}$. Specifically, for each link $(u, v)$ on $p$, we find the BV-T to be used on FlexE-SW $u$ based on the weight $\varpi_{u,v}$ and status of the BV-Ts on $u$ (*Lines* 7-24). Note that, the "first-fit" in *Lines* 9 and 12 means that the first BV-T, which satisfies the corresponding condition, is selected. Finally, *Lines* 29-30 aggregate the CLR schemes of all the flows in **R**, and implement the aggregated CLR schemes in the FlexE-over-EON to recover the flows cost-effectively.




(a) Six-node topology

(b) US Backbone (USB) topology

Fig. 4. Topologies of EON layer with lengths in km marked on links.

**Complexity Analysis:** The time complexity of *Algorithm* 2 can be analyzed as follows. The complexity of the sorting in *Line* 1 is $O(|\mathbf{R}| \cdot \log_2(|\mathbf{R}|))$, where $|\mathbf{R}|$ denotes the number of affected flows in **R**. The for-loop of *Lines* 2-28 runs for $|\mathbf{R}|$ times. In *Lines* 3-4, the complexity of building an AG with *Algorithm* 1 is $O(|V|^2 \cdot |E|)$, and that of calculating $K$ shortest paths in the AG is $O(K \cdot |V|(|\hat{E}_a| + |V| \cdot \log_2(|V|))$ [41], where $|\hat{E}_a|$ denotes the largest number of links in an AG. The complexity of *Lines* 5-25 is $O(K \cdot |V| \cdot |\hat{B}_u|)$, where $|\hat{B}_u|$ is the largest number of BV-Ts on a FlexE-SW. This is because in the worst case, *Lines* 7-24 need to check all the BV-Ts in the FlexE-over-EON to find a CLR scheme $B_p$. Finally, we can obtain the time complexity of *Algorithm* 2 as $O(|\mathbf{R}| \cdot (\log_2(|\mathbf{R}|) + |V|^2 \cdot |E| + K \cdot |V|(|\hat{E}_a| + |V| \cdot \log_2(|V|) + |\hat{B}_u|)))$.

## VI. Performance Evaluations

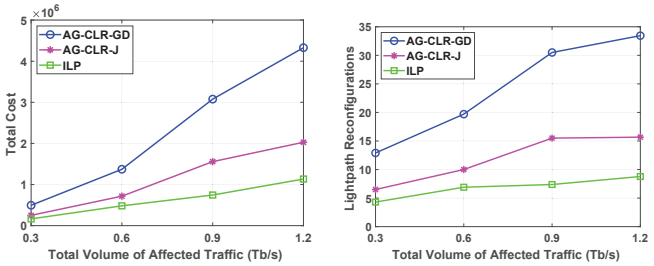
In this section, we discuss the numerical simulations for evaluating the performance of our proposed CLR algorithm.

### A. Simulation Setup

The simulations consider two topologies for the EON layer, *i.e.*, the six-node and US Backbone (USB) topologies in Figs. 4(a) and 4(b), respectively. In the EON layer, we assume that each fiber link can accommodate 358 FS' at most and the bandwidth of each FS is 12.5 GHz. The capacity of an FS depends on the modulation format that it uses, as listed in Table I. Then, each simulation runs as follows. We first generate a set of client flows with a fixed amount of total traffic volume. For each client flow, the source and destination

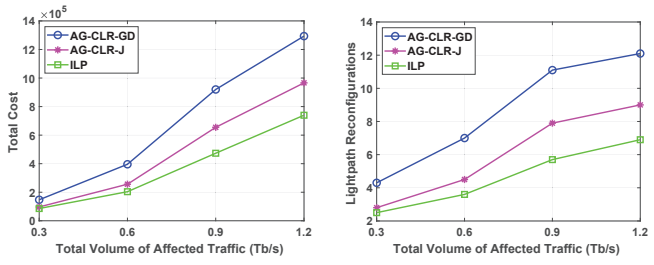
TABLE II
RUNNING TIME OF ALGORITHMS PER AFFECTED FLOW WITH SIX-NODE TOPOLOGY (SECONDS)

| Total Volume of Traffic in **R** (Tb/s) | *Small Flows* | | | *Large Flows* | | |
|---|---|---|---|---|---|---|
| | ILP | AG-CLR-J | AG-CLR-GD | ILP | AG-CLR-J | AG-CLR-GD |
| 0.3 | 0.2376 | 0.0078 | 0.0068 | 0.1817 | 0.0213 | 0.0190 |
| 0.6 | 1.6435 | 0.0063 | 0.0059 | 0.1413 | 0.0114 | 0.0117 |
| 0.9 | 11.3115 | 0.0059 | 0.0053 | 2.3984 | 0.0088 | 0.0082 |
| 1.2 | 15.7711 | 0.0055 | 0.052 | 40.6224 | 0.0092 | 0.0067 |




(a) Total cost (b) Lightpath reconfigurations

Fig. 5. Results of simulations with six-node topology for *Small Flows*.




(a) Total cost (b) Lightpath reconfigurations

Fig. 6. Results of simulations with six-node topology for *Large Flows*.

FlexE-SWs are randomly selected, and its bit-rate is chosen according to one of the two following scenarios [12]

- *Small Flows*: the bit-rate of a flow is randomly selected from $\{10, 40, 25 \cdot \lambda\}$, where the range of $\lambda$ is $[1, 4]$.
- *Large Flows*: the bit-rate of a flow is randomly selected from $\{10, 40, 25 \cdot \lambda\}$, where the range of $\lambda$ is $[5, 8]$.

Here, $\lambda$ is the bit-rate update multiplier of client MAC interfaces [11]. Then, we leverage the multi-hop network planning algorithm developed in [13] to provision all the client flows in the FlexE-over-EON. Here, we assume that each T-Box includes two BV-Ts whose total capacity cannot exceed 400 Gbps. Next, we randomly select a FlexE-SW to fail and obtain the traffic matrix **R** of affected flows. Finally, we apply a CLR algorithm to restore all the affected flows in **R**.

Due to the time complexity of the ILP formulated in Section IV, it can only solve the small-scale CLR problems in the six-node topology. As our proposed *Algorithm* 2 is an AG-based CLR algorithm that considers the affected flows jointly, we refer to it as AG-CLR-J in the following discussions. In addition to ILP and AG-CLR-J, we also introduce a greedy-based benchmark, namely, AG-CLR-GD. Specifically, AG-CLR-GD first sorts the affected flows in **R** in descending order of their bit-rates, and then restores them sequentially in the sorted order by building an AG for each flow and routing the flow in it. The simulations evaluate the algorithms in terms of the total cost of CLR. Here, the dynamic power consumption of each additional FS can be found in Table I, while the static power usages of a BV-T and a T-Box are $w_b = 100$ W and $w_t = 250$ W, respectively. In the optimization, we set the primary objective as to minimize the number of lightpath reconfigurations. Hence, we assume that the unit cost of power consumption is $c_w = 1$ and have $c_r \gg c_w$.

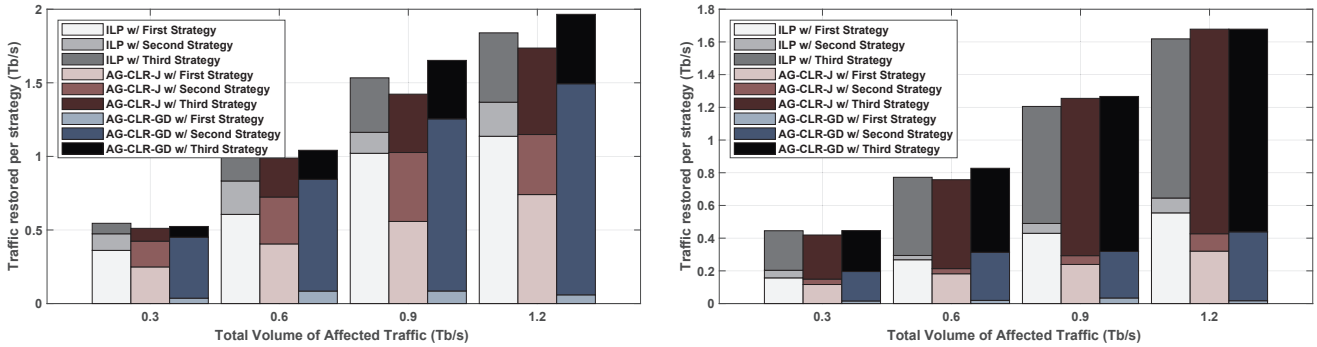
We set $K = 3$ for the $K$ shortest-path routing in AG-CLR-J. To ensure the statistical accuracy of simulation results, we obtain each data point by averaging the results from 10 independent runs. The simulations are performed on a computer with 8 Intel Core i7-6700K CPU @ 4.0 GHz and 64 GB memory, and the software environment is MATLAB 2017b with Gurobi optimization toolbox.

### B. Small-Scale Simulations

We first conduct small-scale simulations in the six-node topology to compare the three algorithms. Figs. 5 and 6 shows the results of the *Small Flows* and *Large Flows* scenarios. For the *Small Flows* scenario, Fig. 5(a) shows that the total cost of the CLR scheme obtained by ILP is always the smallest, followed by that from AG-CLR-J, and AG-CLR-GD performs the worst in terms of the total cost. This confirms the effectiveness of AG-CLR-J. Meanwhile, in Fig. 5(b), we can see that the results on the number of lightpath reconfigurations follow the same trend, which verifies that by considering the affected flows in **R** jointly, AG-CLR-J can effectively reduce the number of lightpath reconfigurations in CLR.

The simulation results of the *Large Flows* scenario are shown in Fig. 6, which follow the same trends as those in Fig. 5. Meanwhile, we notice that the gaps between the results from ILP and AG-CLR-J become smaller. This suggests that AG-CLR-J can better approximate the optimal results from ILP in this scenario. This can be explained as follows. As the *Large Flows* scenario increases the average bit-rate of affected flows, CLR generally needs to invoke more lightpath reconfigurations (*i.e.*, the CLR strategies in Figs. 3(b) and 3(c) will be used more frequently). Therefore, the benefit of AG-CLR-J due to considering the affected flows jointly becomes more obvious.

To further investigate the performance of the algorithms, we plot the distribution of flows recovered with three CLR strategies from each algorithm in Fig. 7. It can be seen that ILP manages to restore the largest volume of affected traffic with the first CLR strategy, followed by AG-CLR-J, while




(a) *Small Flows* (b) *Large Flows*

Fig. 7. Distribution of flows recovered with three CLR strategies (six-node topology).

AG-CLR-GD performs the worst from this perspective. This further explains the benefits of ILP and AG-CLR-J on OPEX saving. Fig. 7 also indicates that with different algorithms, the total volume of affected traffic recovered with the three strategies is different and it is normally greater than the total traffic volume in **R**. This is because an affected flow in **R** can be recovered with a path that goes through multiple lightpaths, and as the affected flow uses a CLR strategy per lightpath, restoring it end-to-end actually takes multiple CLR strategies.

Table II lists the algorithms' running time per affected flow. As expected, ILP is much more time-consuming than AG-CLR-J and AG-CLR-GD, and its running time increases exponentially with the problem scale. Hence, although ILP can plan CLR schemes the best, it has difficulty to realize timely CLR. Meanwhile, both AG-CLR-J and AG-CLR-GD run very fast to plan the CLR scheme of an affected flow within 30 msec, which verifies that they can quickly resolve packet layer outages in the FlexE-over-EON to satisfy the service-level agreements (SLAs) for the emerging network services that are sensitive to latency and reliability. Note that, existing failure recovery techniques like the Fast Reroute [42] can reduce the restoration time of a flow to less than 50 msec. Fast Reroute is a protection technology, which means that a backup path needs to be pre-established for each flow and path switching will be performed during a network failure. However, our CLR scheme is a restoration technology, which only calculates the restoration path for each affected flow after a switch outage actually has happened. Therefore, since a restoration technology will not pre-establish any backup paths, the restoration time of our CLR scheme will be longer than that of Fast Reroute, but our CLR scheme does not need to reserve any backup resources in an FlexE-over-EON and thus makes its resource utilization more efficient than Fast Route.

### C. Large-Scale Simulations

Next, we consider the USB topology for large-scale simulations. This time, ILP becomes intractable and thus the simulations only consider AG-CLR-J and AG-CLR-GD. Figs. 8 and 9 respectively show the results of the *Small Flows* and *Large Flows* scenarios, and Fig. 10 illustrates the distribution of flows recovered with three CLR strategies from each algorithm. Here, the trends in Figs. 8(a) and 8(b), 9(a) and 9(b), and 10 are similar to those in Figs. 5, 6, and 7, respectively, and confirm the advantages of AG-CLR-J over AG-CLR-GD.

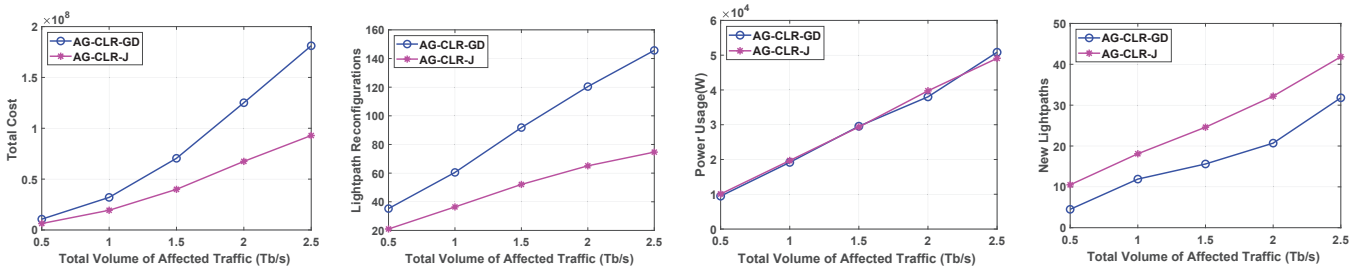
Figs. 8(c) and 9(c) compare the extra power consumption introduced by CLR, which indicate that the extra power usages of the CLR schemes from the two algorithms are similar. This suggests that AG-CLR-J reduces the total cost of CLR mainly by invoking much less lightpath reconfigurations than AG-CLR-GD, which coincide with our setting of the primary optimization objective (*i.e.*, to minimize lightpath reconfigurations). Figs. 8(d) and 9(d) show the results on the number of new lightpaths built in CLR. It is interesting to notice that AG-CLR-J actually establishes more new lightpaths than AG-CLR-GD. This is because AG-CLR-J considers affected flows jointly. In other words, even though AG-CLR-J sets up slightly more new lightpaths in Figs. 8(d) and 9(d), the total lightpath reconfigurations invoked by it are still significantly fewer than those by AG-CLR-GD (as shown in Figs. 8(b) and 9(b)). This is the reason why AG-CLR-J can achieve lower total costs in Figs. 8(a) and 9(a). Table III lists the algorithms' running time per affected flow. We can see that AG-CLR-J and AG-CLR-GD have similar running time and they both can still plan the CLR scheme of an affected flow within 30 msec.

## VII. Conclusion

This paper studied the problem of CLR in a FlexE-over-EON that uses the FlexE-aware architecture. Specifically, we considered the situation where an outage happened on one FlexE-SW to bring it offline temporarily and then the affected client flows need to be recovered quickly and proactively. In view of the flexibility of FlexE-over-EON, three CLR strategies were first proposed to restore the affected flows. Then, to minimize the additional OPEX caused by CLR, we formulated an ILP model and designed an AG-based CLR algorithm. Through extensive simulations, we evaluated the proposed CLR algorithms. The results confirmed that our proposed AG-based algorithm can approximate the optimal results from the ILP and outperform a greedy-based benchmark significantly, thereby effectively reducing the additional OPEX of CLR.

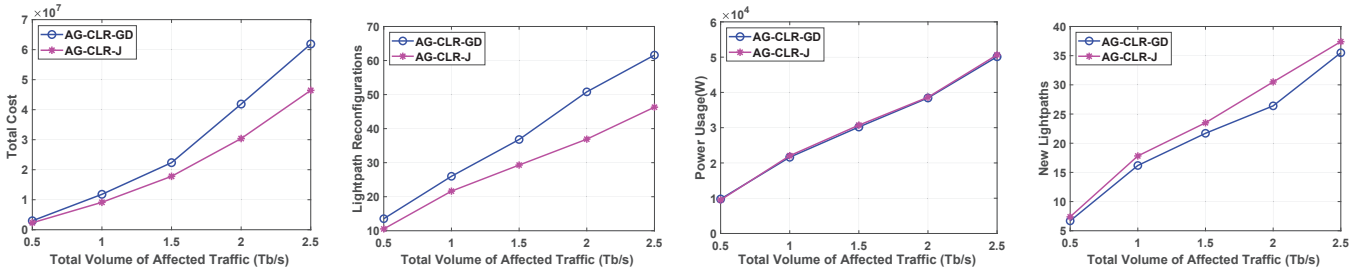
## Acknowledgments

This work was supported in part by the NSFC project 61871357 and Fundamental Funds for Central Universities (WK3500000006).



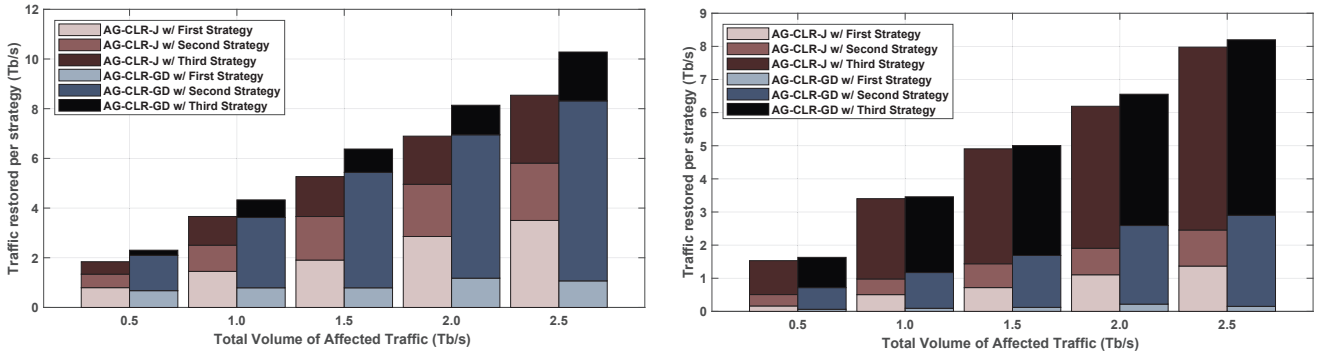

(a) Total cost (b) Lightpath reconfigurations (c) Power consumption (d) New lightpaths

Fig. 8. Results of simulations with USB topology for *Small Flows*.




(a) Total cost (b) Lightpath reconfigurations (c) Power consumption (d) New lightpaths

Fig. 9. Results of simulations with USB topology for *Large Flows*.




(a) *Small Flows* (b) *Large Flows*

Fig. 10. Distribution of flows recovered with three CLR strategies (USB topology).

## REFERENCES


[1] Cisco Visual Networking Index, 2017-2022. [Online]. Available: https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white-paper-c11-741490.html.

[2] P. Lu *et al.*, "Highly efficient data migration and backup for Big Data applications in elastic optical inter-data-center networks," *IEEE Netw.*, vol. 29, pp. 36–42, Sept./Oct. 2015.

[3] K. Han *et al.*, "Application-driven end-to-end slicing: When wireless network virtualization orchestrates with NFV-based mobile edge computing," *IEEE Access*, vol. 6, pp. 26 567–26 577, 2018.

[4] Y. Seol *et al.*, "Timely survey of time-sensitive networking: Past and future directions," *IEEE Access*, vol. 9, pp. 142 506–142 527, 2021.

[5] "Flex Ethernet (FlexE)." [Online]. Available: https://www.oiforum.com/technical-work/hot-topics/flex-ethernet-flexe-2/.

[6] M. Zhang, "Flex Ethernet technology and application in 5G mobile transport network," *China Commun.*, vol. 18, pp. 250–258, Feb. 2021.

[7] N. Huin *et al.*, "Hard-isolation for network slicing," in *Proc. of INFOCOM WKSHPS 2019*, pp. 955–956, May 2019.

[8] L. Gong and Z. Zhu, "Virtual optical network embedding (VONE) over elastic optical networks," *J. Lightw. Technol.*, vol. 32, pp. 450–460, Feb. 2014.

[9] L. Gong, H. Jiang, Y. Wang, and Z. Zhu, "Novel location-constrained virtual network embedding (LC-VNE) algorithms towards integrated node and link mapping," *IEEE/ACM Trans. Netw.*, vol. 24, pp. 3648–3661, Dec. 2016.

[10] J. Liu *et al.*, "On dynamic service function chain deployment and readjustment," *IEEE Trans. Netw. Serv. Manag.*, vol. 14, pp. 543–553, Sept. 2017.

[11] A. Eira, A. Pereira, J. Pires, and J. Pedro, "On the efficiency of flexible Ethernet client architectures in optical transport networks," *J. Opt. Commun. Netw.*, vol. 10, pp. A133–A143, Jan. 2018.

[12] W. Lu *et al.*, "How much can flexible Ethernet and elastic optical networking benefit mutually?" in *Proc. of ICC 2019*, pp. 1–6, May 2019.

[13] H. Liang, N. L. S. da Fonseca, and Z. Zhu, "On the cross-layer network planning for flexible Ethernet over elastic optical networks," *IEEE Trans. Netw. Serv. Manag.*, vol. 18, pp. 3691–3705, Sept. 2021.

[14] D. Koulougli, K. Nguyen, and M. Cheriet, "Efficient routing using flexible Ethernet in multi-layer multi-domain networks," *J. Lightw. Technol.*, vol. 39, pp. 1925–1936, Apr. 2021.

[15] Z. Zhu, W. Lu, L. Zhang, and N. Ansari, "Dynamic service provisioning in elastic optical networks with hybrid single-/multi-path routing," *J. Lightw. Technol.*, vol. 31, pp. 15–22, Jan. 2013.

[16] L. Gong *et al.*, "Efficient resource allocation for all-optical multicasting over spectrum-sliced elastic optical networks," *J. Opt. Commun. Netw.*, vol. 5, pp. 836–847, Aug. 2013.

[17] Y. Yin *et al.*, "Spectral and spatial 2D fragmentation-aware routing and spectrum assignment algorithms in elastic optical networks," *J. Opt.*

TABLE III
RUNNING TIME OF ALGORITHMS PER AFFECTED FLOW WITH USB TOPOLOGY (SECONDS)

| Total Volume of Traffic in **R** (Tb/s) | *Small Flows* | | *Large Flows* | |
|---|---|---|---|---|
| | AG-CLR-J | AG-CLR-GD | AG-CLR-J | AG-CLR-GD |
| 0.5 | 0.0275 | 0.0270 | 0.0275 | 0.0271 |
| 1.0 | 0.0275 | 0.0268 | 0.0271 | 0.0267 |
| 1.5 | 0.0278 | 0.0275 | 0.0270 | 0.0267 |
| 2.0 | 0.0280 | 0.0278 | 0.0264 | 0.0261 |
| 2.5 | 0.0277 | 0.0274 | 0.0271 | 0.0266 |

*Commun. Netw.*, vol. 5, pp. A100–A106, Oct. 2013.
[18] M. Zhang, C. You, H. Jiang, and Z. Zhu, "Dynamic and adaptive bandwidth defragmentation in spectrum-sliced elastic optical networks with time-varying traffic," *J. Lightw. Technol.*, vol. 32, pp. 1014–1023, Mar. 2014.
[19] M. Jinno *et al.*, "Multiflow optical transponder for efficient multilayer optical networking," *IEEE Commun. Mag.*, vol. 50, pp. 56–65, May 2012.
[20] N. Sambo *et al.*, "Next generation sliceable bandwidth variable transponders," *IEEE Commun. Mag.*, vol. 53, pp. 163–171, Feb. 2015.
[21] Z. Zhu *et al.*, "Demonstration of cooperative resource allocation in an OpenFlow-controlled multidomain and multinational SD-EON testbed," *J. Lightw. Technol.*, vol. 33, pp. 1508–1514, Apr. 2015.
[22] M. Liu, M. Tornatore, and B. Mukherjee, "Survivable traffic grooming in elastic optical networks - shared protection," *J. Lightw. Technol.*, vol. 31, pp. 903–909, Mar. 2013.
[23] F. Ji *et al.*, "Dynamic p-cycle protection in spectrum-sliced elastic optical networks," *J. Lightw. Technol.*, vol. 32, pp. 1190–1199, Mar. 2014.
[24] X. Chen, F. Ji, and Z. Zhu, "Service availability oriented p-cycle protection design in elastic optical networks," *J. Opt. Commun. Netw.*, vol. 6, pp. 901–910, Oct. 2014.
[25] X. Chen *et al.*, "Flexible availability-aware differentiated protection in software-defined elastic optical networks," *J. Lightw. Technol.*, vol. 33, pp. 3872–3882, Sept. 2015.
[26] X. Chen, S. Zhu, L. Jiang, and Z. Zhu, "On spectrum efficient failure-independent path protection p-cycle design in elastic optical networks," *J. Lightw. Technol.*, vol. 33, pp. 3719–3729, Sept. 2015.
[27] R. Govindan *et al.*, "Evolve or die: High-availability design principles drawn from Google's network infrastructure," in *Proc. of ACM SIGCOMM 2016*, pp. 58–72, Aug. 2016.
[28] S. Liu, W. Lu, and Z. Zhu, "On the cross-layer orchestration to address IP router outages with cost-efficient multilayer restoration in IP-over-EONs," *J. Opt. Commun. Netw.*, vol. 10, pp. A122–A132, Jan. 2018.
[29] K. Christodoulopoulos, I. Tomkos, and E. Varvarigos, "Elastic bandwidth allocation in flexible OFDM-based optical networks," *J. Lightw. Technol.*, vol. 29, pp. 1354–1366, May 2011.
[30] L. Gong, X. Zhou, W. Lu, and Z. Zhu, "A two-population based evolutionary approach for optimizing routing, modulation and spectrum assignments (RMSA) in O-OFDM networks," *IEEE Commun. Lett.*, vol. 16, pp. 1520–1523, Sept. 2012.
[31] S. Zhang *et al.*, "Evolving traffic grooming in multi-layer flexible-grid optical networks with software-defined elasticity," *J. Lightw. Technol.*, vol. 32, pp. 2905–2914, Aug. 2014.
[32] S. Liu *et al.*, "DL-assisted cross-layer orchestration in software-defined IP-over-EONs: From algorithm design to system prototype," *J. Lightw. Technol.*, vol. 37, pp. 4426–4438, Sept. 2019.
[33] P. Papanikolaou, K. Christodoulopoulos, and E. Varvarigos, "Joint multilayer survivability techniques for IP-over-elastic-optical-networks," *J. Opt. Commun. Netw.*, vol. 9, pp. A85–A98, Jan. 2017.
[34] W. Lu, X. Yin, X. Cheng, and Z. Zhu, "On cost-efficient integrated multilayer protection planning in IP-over-EONs," *J. Lightw. Technol.*, vol. 36, pp. 2037–2048, May 2018.
[35] S. Trowbridge, "Ethernet and OTN: 400G and beyond," in *Proc. of OFC 2015*, pp. 1–3, Mar. 2015.
[36] T. Hofmeister, V. Vusirikala, and B. Koley, "How can flexibility on the line side best be exploited on the client side?" in *Proc. of OFC 2016*, pp. 1–3, Mar. 2016.
[37] P. Zhu, J. Cui, and Y. Ji, "Universal hash based built-in secure transport in FlexE over WDM networks," *J. Lightw. Technol.*, vol. 39, pp. 5680–5690, Sept. 2021.
[38] D. Koulougli, K. Nguyen, and M. Cheriet, "Flexible Ethernet traffic restoration in multi-layer multi-domain networks," in *Proc. of ICC 2021*, pp. 1–6, Jun. 2021.
[39] A. Bocoi *et al.*, "Reach-dependent capacity in optical networks enabled by OFDM," in *Proc. of OFC 2009*, pp. 1–3, Mar. 2009.
[40] W. Lu and Z. Zhu, "Dynamic service provisioning of advance reservation requests in elastic optical networks," *J. Lightw. Technol.*, vol. 31, pp. 1621–1627, May 2013.
[41] J. Yen, "Finding the k-shortest loopless paths in a network," *Manag. Sci.*, vol. 17, pp. 712–716, Jul. 1971.
[42] A. Raj and O. Ibe, "A survey of IP and Multiprotocol Label Switching fast reroute schemes," *Comput. Netw.*, vol. 51, pp. 1882–1907, Jun. 2007.