Exploiting SDN Principles for Extremely Fast Restoration in Elastic Optical Datacenter Networks

Yu Xiong*†, Yuanyuan Li*, Xiancun Dong*, Yupeng Gao*, Gerorge N. Rouskas†
*Key Laboratory of Optical Communication and Networks,
Chongqing University of Posts and Telecommunications, Chongqing, 400065, China
†Operations Research and Department of Computer Science,
North Carolina State University, Raleigh, NC 27695-8206, USA
Email: cqxiongyu@foxmail.com, liyuanyuancq@163.com, dongxiancun@foxmail.com,
gaoyupengcqupt@foxmail.com, rouskas@ncsu.edu

***Abstract*—In order to find out the tradeoff between recovery time and resource overhead when a link failure occurs, we propose a software-defined based fast restoration scheme for elastic optical datacenter networks, called precomputation based restoration path (P-RP). By extending the controller functionality and OpenFlow protocol in software defined networking (SDN) technology, we establish a novel elastic optical inter-datacenter network architecture which can quickly and accurately converge the network state information with a global view for failure recovery. Based on this architecture, the P-RP performs path computation and bandwidth resource selection for restoration before a failure occurs. The experimental results demonstrate that the proposed scheme can achieve fast recovery with low blocking probability while maintaining high spectrum efficiency. Compared with existing restoration schemes, average recovery time is improved by up to 28%.**

Keywords*—*elastic optical network*; *inter-datacenter*; *restoration*; *software defined networking

I. Introduction

With the emergence of big data, cloud computing and high-definition multimedia video, IP services in datacenter networks are growing at an amazing rate [1]. Due to the diversity and point-to-content characteristics of data center services, they utilize optical network transport to provide flexibility, speed and reliability [2]. The elastic optical network (EON) is based on orthogonal frequency division multiplexing technology. It supports channels operating at heterogeneous line rates by allocating spectral resources on demand for a variety of connection demands in a flexible and dynamic manner [3], and hence it is suitable for datacenter network scenarios [4]. In interconnecting data centers, EONs enable communication between large-scale servers, and form the foundation of the network storage and computation. Due to the high transmission capacity of single fiber in EON, a single link failure will result in high data loss. Obviously, the survivability of elastic optical networks interconnecting data centers is an important and meaningful issue.

A service-aware protection with bandwidth squeezing to protect against disaster in elastic optical inter-datacenter network was proposed in [5]. This mechanism improves the spectrum resource utilization, while guaranteeing the quality of service (QoS) requirements of different services. Another study takes into consideration the anycast characteristics of datacenter applications to propose a bandwidth-adaptable protection with content connectivity against disasters in the context of elastic optical inter-datacenter networks [6]. This approach not only ensures content connectivity, but also reduces the blocking rate. These studies were designed for traditional optical networks with distributed control, resulting in lower flexibility and the higher network cost. Furthermore, data center services show high-burstiness and high-bandwidth characteristics, and hence require the control plane of optical networks to respond rapidly to service requests; this requirement brings great scheduling pressure to the network. Recently, as a promising centralized control architecture, the software defined networking (SDN) enabled by OpenFlow protocol supports programmability of data center and network functionalities, which can provide maximum flexibility for operators and make an unified control over various resources for the joint optimization of functionalities and services in a global view [7]. In [8], an OpenFlow-based software defined elastic optical network architecture is proposed for data center service protection. Further, a multipath protection scheme is figured based on the importance level of the service to ensure the QoS at the lower cost. However, the protection scheme needs to reserve resources for the services on backup paths, while the restoration scheme is not needed. So the restoration scheme can effectively improve the resources utilization rate. Therefore, some restoration schemes are proposed for elastic optical networks based on SDN framework [9], [10]. By centralized control method of SDN, the controller quickly perceives the failure information; then, rerouting procedure is used to recalculate the route for services affected by failures; and finally the configuration commands are issued in parallel from the controller to complete the service restoration process. However, these schemes do not consider the problem that the calculation time of the rerouting accounts for a larger proportion of the total recovery time.

In this paper, we propose a new precomputation based restoration path (P-RP) scheme to provision extremely fast and efficient restoration in elastic optical inter-datacenter networks. Meanwhile, we leverage the distance-adaptive routing and spectrum assignment algorithm when P-RP scheme implements restoration path computation and bandwidth resource allocation before a failure occurs. Furthermore, the software-defined elastic optical inter-datacenter network architecture is redesigned by extending OpenFlow protocol and functionality of controller for the datacenter application requests. Finally, the proposed scheme is performed and simulated on software-defined elastic optical inter-datacenter network testbed. The rest of this paper is organized as follows. In section, the software-defined elastic optical inter-datacenter network architecture is designed with OpenFlow protocol extension and functionality module of controller. In section III, P-RP scheme is proposed for data center services under this control architecture. numeric results and analysis are given in section IV. Finally, Section V concludes the paper and presents some future work.

II. Network Architecture

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 centralized controlled by the controller with extended OpenFlow protocol (OFP). In order to realize the control functionality of the controller, the network nodes on the underlying data plane should be composed of OFP agent and OpenFlow-enabled bandwidth variable optical switch (OF-BVOS). The OFP agent receives flow table messages from controller, and then translates it into the logical language, which the underlying hardware devices can understand, and then control the cross connection process of the underlying OF-BVOS. In 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 restoration scheme associated with the optical network resources. Finally, configuration commands are issued to complete the restoration in parallel by the controller through the extended OFP.

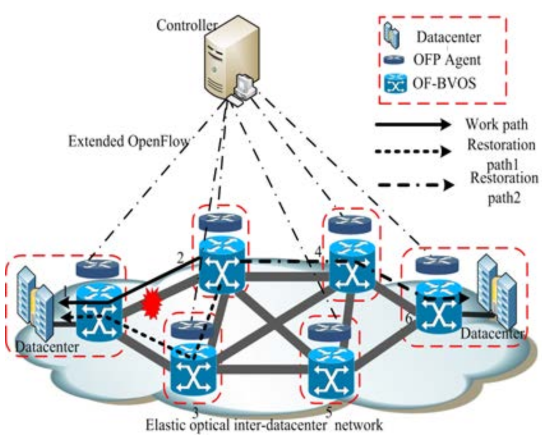

Fig. 1: Network architecture based on different interest communities.

In order to support the functionalities of the above architecture, the OpenFlow protocol and controller need be extended. The datacenter ID is added to the OFP on the basis of the reference [11]; then the functionality module of the controller is extended as in Fig.2. The OpenFlow controller consists of seven modules, i.e., network abstraction, path computing entity (PCE) and plug-in, data base modules, forward Open API, precomputation and restoration, failure detection and Web management. The responsibilities and interactions among the functional modules are provided below.

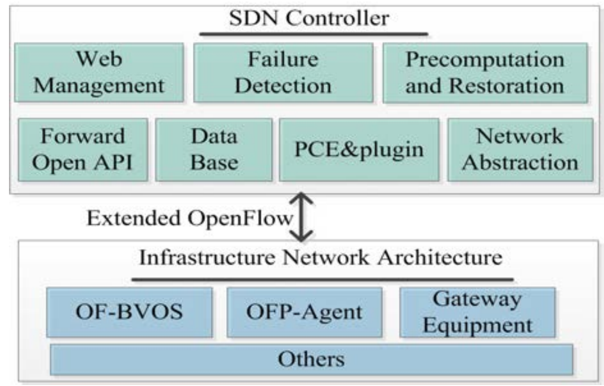

Fig. 2: The functional model of controller in software defined elastic optical network architecture.

The network abstraction 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 extended OFP. In case of link failures, the failure detection module discovers them and delivers such failure information to the precomputation and restoration module. Before the link failure occurs in elastic optical network, the precomputation and restoration module decide to apply precomputation scheme associated with the optical network resources. The PCE module can calculate the working and restoration lightpaths interworking with the OFP agent of the OF-BVOS, where the various computation strategies are alternative as a plug-in. The information of the path is conserved into data base management and updated the results of paths.

III. Precomputation based Restoration Path Scheme (P-RP)

The precomputation restoration strategy is proposed to reduce the high weight of calculation time of the rerouting procedure during restoration scheme [12]. However, in this distributed network, the 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. Precomputation before a Failure Occurs

In the elastic optical inter-datacenter networks, we leverage SDN technology to precompute the restoration path, which just occupied little CPU computing resources through an extra opening thread in the controller. PR-P scheme can achieve for the precomputation based restoration path without consuming the transmission resources. 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. So the performance for P-RP based on SDN control technology is also superior. The precomputation restoration strategy, firstly, precompute 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 the service really interrupts, the controller can directly use the precomputation recovery scheme that stored in the controller and allocate the network resources for the interruption of service. In this paper, we only consider that the availability of the required resources is ensured. Thus, the restoration route can be successfully set up.

While in datacenter networks, most requests call for application services, such as cloud computation and video, which are stored in datacenters. Those types of requests are anycast requests, where the destination node is not fixed, and any datacenter with the required content service can be selected as the destination node. Here, D is defined as set of datacenters, $|D|$ 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 Dijkstra algorithm is used when establishing connection requests[13]. Any service r can find $|D|$ shortest paths from source node to each datacenter with Dijkstra algorithm. Then, the link cost functionality is defined to select an optimal backup restoration route from $|D|$ paths, as shown in formula (1). According to formula (1), the path with the minimum link cost functionality is selected as the backup restoration path. For a request $\mathrm{r}(s, d, \Phi)$, where s, d and Φ are respectively source node, destination node and bandwidth of working path, the required number of FSs for a certain modulation format can be easily found according to formula (2) and (3). We only consider three modulation format such as BPSK, QPSK and 8QAM. The parameters of modulation format are shown in Table .

$$C = H_{\mathrm{p}_{s,d_i}} \times m_{sb} \tag{1}$$

Where $H_{p_{s,d_i}}$ is the hop of p_{s,d_i}; m_{sb} is the number of spectrum blocks.

$$M_i = Mod(\sum_e L_e), e \in R_{s,d,i} \tag{2}$$

$$n_{\mathrm{i}} = \frac{\Phi}{B \cdot \log_2 M_{\mathrm{i}}} \tag{3}$$

Where $Mod(\cdot)$ returns the highest modulation level that transmission distance can support; n_i is the required FSs. L_e is the length of each link. B is bandwidth of each sub-carrier slot.

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 through $PORT\_STATUS$ messages.The

TABLE I: Parameters of Modulation Formats[14]

Modulation Format	Symbol rate	Bit per symbol
BPSK	B/2n(const)	1
QPSK	B/2n(const)	2
8QAM	B/2n(const)	3

controller searches for the precomputation restoration route information, which is 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 are remained in the flow table of source node, and the packet will be forwarded along the remained 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. The interworking procedure for restoration based on precomputation, when link failure occurs, is shown in Fig.3. 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 scheme stored in controller. τ_i' represents the time used by controller to send $FLOW\_MOD$ messages labeled with ”DELETE” to the nodes of the affected paths. Δt is the time interval between the two controller issues the message labeled with ”ADD” and the message labeled with ”DELETE”. Generally, $\Delta t \geq 0$.

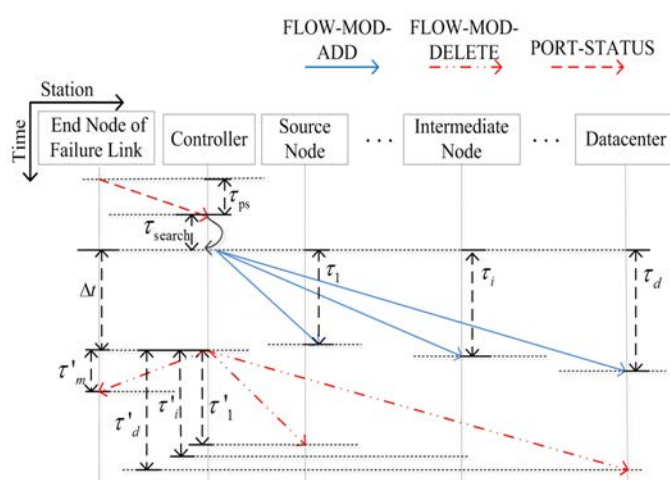

Fig. 3: Precomputation restoration strategy based on SDN timeline.

Hence, the restoration time of the whole restoration scheme based on precomputation is shown in formula (4).

$$\mathrm{t}_{pc} = \tau_{ps} + \tau_{search} + Max[(\Delta t + \underset{i\in[1.n]}{Max} {\tau'}_i),\ \underset{i\in[1,n]}{Max}\ \tau_i] \tag{4}$$

Traditional restoration scheme has larger delay, which is mainly produced by the process of computing bandwidth, center frequency and the cross connection process between

OF-BVOSs. What's more, the procedure of computing bandwidth and center frequency account for large proportion of the restoration time. The restoration scheme based on precomputation can solve the problem of large delay produced by bandwidth and center frequency computation.

IV. Simulation Analysis

In this section, in order to demonstrate that the P-RP restoration scheme can meet the network survivability requirements, we leverage the Mininet+Floodlight simulation tool in the network topology of NSFNet (14 nodes, 21 links), as shown in Fig.4. NSFNet has datacenters at nodes 0, 5 and 8 [6], [15]. And we evaluate the performance of P-RP scheme under heavy traffic load scenario and compare it with the traditional shared protection based SDN(SDN-TSP) and fast restoration based SDN (SDN-ind)[9] in the software defined elastic optical inter-datacenter network, but we only consider single-link failure of SDN-ind here. The described schemes are evaluated in C++ based Linux GCC tool on a computer. Assuming that there is one pair of bi-directional fiber on each link, and the available spectrum width of each fiber is set to 4000GHz with the width of 12.5GHz. In this simulation, the flow requests to datacenter nodes are established with bandwidth randomly distributed from 12.5 Gbps to 100 Gbps, which arrive at network following a Poisson process with the arrival rate μ, and the traffic load is λ/μ (Erlang) [8].The time parameters settings are also the same as that of reference [9].

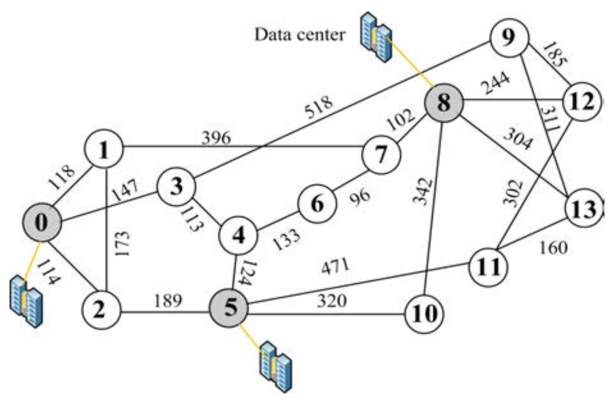

Fig. 4: NSF network topology

In this paper, the three performance indexes are considered, which are average recovery time (ART), blocking probability (BP) and resource occupation rate (ROR). ART refers to the link failure recovery time, which consists of the failure detection time and provisioning time of a new path; BP is the percentage of the number of rejected services accounted for all the arrival services in the network; ROR reflects the percentage of occupied working frequency slots resources to the total number of network frequency slots resources.

Fig.5 illustrates the capture procedure of the OpenFlow messages, which comprises controller-switch, asynchronous and modify-state messages. The three schemes are evaluated using the same routing and spectrum assignment strategy

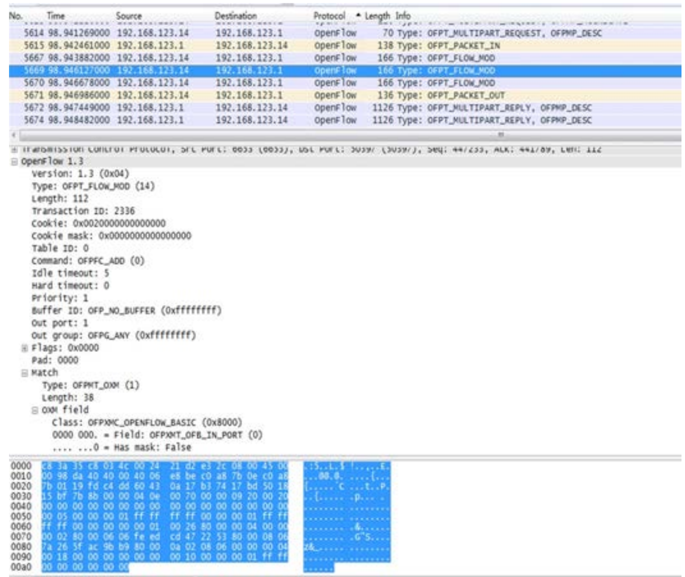

Fig. 5: The capture of OpenFlow messages

during both provisioning and restoration phases. Obviously, the restoration time of SDN-ind scheme is longer than that of the SDN-TSP and P-RP schemes, which can be seen in Table II. We built a testbed with the topology as shown in Fig.1. to evaluate the time performance of proposed mechanisms. We first set up three paths,as shown in Table . We built a testbed with the topology as shown in Fig.1. to evaluate the time performance of proposed mechanisms. We first set up three paths, as shown in Table . We break the link between OF-BVOS1 and OF-BVOS2, and the protection path and restoration path are respectively 2-3-1 and 2-4-6. Each simulation data is taken 100 times average. The recovery time of protection scheme mainly depends on the failure detection time and the deletion time of working flow in source node, while the time of restoration scheme depends on the failure detection time, computation time and provisioning time of a new path. Compared with SDN-ind scheme using the rerouting method in reference [9], since the proposed P-RP uses precomputation restoration method to effectively reduce the complicated computation time, it greatly improves the recovery time after the failure occurs. So the reason for the difference is the latency of computation time of a new path. As shown in Table II, the ART of P-RP scheme can improve by 28% than SDN-ind.

As shown in Fig.6, SDN-TSP has the highest BP due to the resource consumption of protection path. On the contrary, P-RP has the lowest BP. Specially, P-RP reduces 57.25% of BP to TSP scheme under 700 Erlang traffic load. That is because the protection algorithm needs to reserve protection resources on backup path for each service in advance, which occupy many spectrum resources of network. And P-RP algorithm that use distance adaptive modulation scheme can transmit the same services with less spectrum resource. And under low traffic load, the blocking probabilities of P-RP and SDN-ind is lower than 10e-2. That is because the network resource

TABLE II: Average recovery time

Recovery Type	Working Path	Protection Path	Restoration path	Recovery Time
SDN-TSP	2-1	2-3-1	Null	<40ms
SDN-ind	2-1	Null	2-3-1	∼85ms
P-RP	2-1	Null	2-4-6	∼61ms

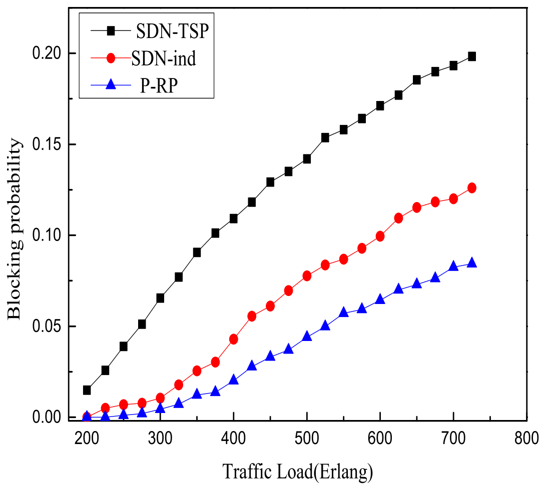

Fig. 6: BP vs. Traffic load

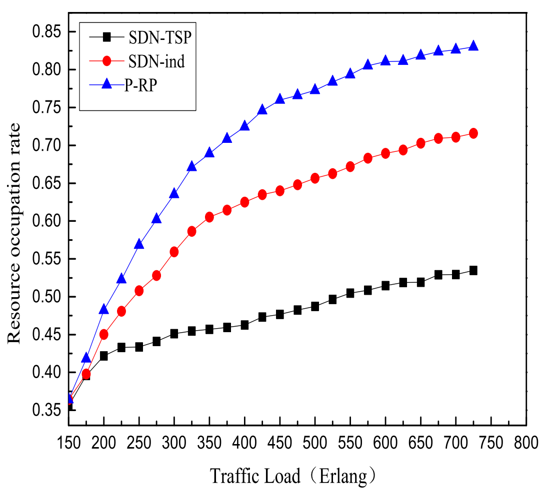

Fig. 7: ROR vs. Traffic load

is sufficient under lower load, so it is with smaller blocking probabilities. The ROR of P-RP is the highest in Fig.7. This is because the P-RP scheme with lower BP accommodates more connection requests, leading to higher resource occupation rate. The BP of P-RP is higher than that of SDN-ind. That is because P-RP algorithm with distance adaptive modulation approach can use the same spectrum to transmit more high-bandwidth services successfully. In summary, although the recovery time of P-RP is a little higher than the SDN-TSP scheme, P-RP scheme can achieve better BP and higher ROR compared to SDN-TSP. Therefore, the proposed P-RP scheme can achieve better network performance in elastic optical inter-datacenter networks based on SDN.

V. CONCLUSIONS

In this paper, we consider the problem of precomputed global rerouting for fault management in software defined elastic optical inter-datacenter networks. We also applied optimization technique to precompute restoration path for data center services, and finally we have successfully implemented OF-based failure recovery mechanisms in software defined elastic optical inter-datacenter networking architecture. The proposed P-RP scheme provides the restoration path and spectrum resources messages before link failure. The feasibility and performances are evaluated through the simulation in terms of blocking probability, average recovery latency and resource occupation rate under different loads. Simulation results indicate that our failure recovery mechanisms can provide reliable and fast restoration under extended OF-based control plane framework. Compared with the TSP and SDN-ind schemes, the proposed P-RP scheme can extremely fast restore the interrupting services using the precomputed restoration path and bandwidth at the cost of the lower network resource overhead.

Future interesting works will be focused on the influence on dynamic parameters of the scheme, which possibly makes for further improving the resource utilization. Simultaneously, the problem about the actual availability of the pre-computed resources before failures occurring should be further considered. Also, the scheme under multiple failures scenarios, the service-aware recovery and experimental demonstration on our OpenFlow-based SDN testbed will be researched in our future works.

REFERENCES

[1] Cisco Visual Networking Index: Forecast and Methodology, 2014-2019, Cisco Syst, San Jose, CA, USA, Feb, 2015.

[2] W. Zhang, H. Wang, K. Bergman, “Next-generation optically-interconnected high-performance data centers,” J. Lightw. Technol, vol. 30, no. 24, pp. 3836-3844, Dec. 2012.

[3] S. Talebi, F. Alam, I. Katib, M. Khamis, Reda. Salama, G. N. Rouskas, “Spectrum management techniques for elastic optical networks: A survey,” Optical Switching and Networking, vol. 13, pp. 3836-3844, Dec. 2014.

[4] M. Jinno, H. Takara, B. Kozicki, Y. Tsukishima, Y. Sone, S. Matsuoka, “Spectrum-efficient and scalable elastic optical path network: architecture, benefits, and enabling technologies,” IEEE Communications Magazine, vol. 47, no. 11, pp. 66-73, Nov. 2014.

[5] Z. Jin, J. Zhang ,Y. Zhao, et al., “Service-Aware Protection with Bandwidth Squeezing against Disaster in Elastic Optical Datacenter Networks,”. in Proc. ACPC2014, shanghai, China, Paper AF3E. 8.

[6] C. Ma, J. Zhang, Y. Zhao, et al., “Bandwidth-adaptability protection with content connectivity against disaster in elastic optical datacenter networks,” Photonic Network Communications, vol. 30, no. 2, pp. 309-320, Jun. 2015.

[7] L. Liu, W. R. Peng, R. Casellas, et al., “Experimental demonstration of OpenFlow-based dynamic restoration in elastic optical networks on GENI testbed,” European Conference on Optical Communications (ECOC), Cannes, France, 2014, pp. 1-3.

[8] H. Yang, L. Cheng, J. Yuan, J. Zhang, Y. L. Zhao, Y. Lee, “Multipath protection for datacenter services in OpenFlow-based software defined elastic optical networks,” Optical Fiber Technology, vol. 23, pp. 108-118, 2015.

[9] A. Giorgetti, F. Paolucci, F. Cugini, et al., “Fast restoration in SDN-based flexible optical networks,” In OFC/NFEOC, California, America, Paper Th3B2, 2014.

[10] F. Paolucci, A. Castro, F. Cugini, et al., “Multipath restoration and bitrate squeezing in SDN-based elastic optical networks [Invited],” Photonic network communications, vol. 28, no. 1, pp. 45-57, 2015.

[11] X. Chen, M. Tornatore, S. Zhu, et al., “Flexible availability-aware differentiated protection in software-defined elastic optical networks,” J. Lightw. Technol, vol. 33, no. 18, pp. 3872-3882, Jul. 2015.

[12] A. Giorgetti, N. Sambo, I. Cerutti, et al., “Label preference schemes for lightpath provisioning and restoration in distributed GMPLS networks,” J. Lightw. Technol, vol. 27, no. 6, pp. 688-697, Jul. 2009.

[13] S. Talebi, I. Katib, G. N. Rouskas, “Offline Distance-Adaptive Routing and Spectrum Assignment (DA-RSA) in Rings,” in Proc. IEEE GLOBECOM, San Diego, USA, 2015:1-6.

[14] M. Jinno, T. Ohara, Y. Sone, et al., “Elastic and adaptive optical networks: possible adoption scenarios and future standardization aspects,” IEEE Communications Magazine, vol. 49, no. 10, pp. 164-172, Oct. 2011.

[15] Habib M F, Tornatore M, De Leenheer M, et al. “A disaster-resilient multi-content optical datacenter network architecture,” In Proc. IEEE ICTON, Stockholm, Sweden, pp. 1-4, 2011.