Cross-Layer Restoration to Resolve Packet Layer Outages in P2MP-TRX-based WSONs

Meihan Wu, Yuxiao Zhang, Ruoxing Li, Qian Lv, Sijia Li, and Zuqing Zhu†
School of Information Science and Technology, University of Science and Technology of China, Hefei, China
†Email: {zqzhu}@ieee.org

***Abstract*—We design three strategies for recovering coherent point-to-multipoint transceiver (P2MP-TRX) based wavelength-switched optical networks (WSONs) from packet layer outages, and propose an efficient algorithm to leverage them for realizing cost-effective cross-layer restoration (CLR).**

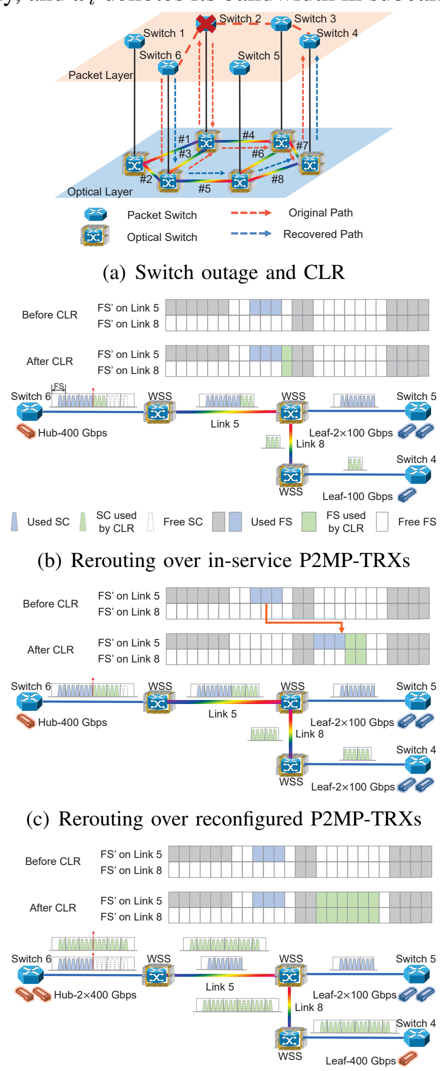
I. Introduction

The invention of coherent point-to-multipoint transceivers (P2MP-TRXs) [1] can make metro-aggregation networks more cost-effectively adapt to the increase of hub-and-spoke (H&S) traffic [2–4]. Meanwhile, previous study has confirmed that a reasonable combination of P2MP-TRXs and wavelength-switched optical networks (WSONs) is promising in terms of scalability, reconfigurability and spectral efficiency [5]. Hence, it is of great importance to study P2MP-TRX-based WSONs, especially for the network survivability as P2MP-TRXs can increase transmission capacity effectively.

The survivability of WSON can be affected by failures in optical and packet layers, and those in packet layer usually happen more frequently [6, 7]. Previously, people have studied the multilayer protection [8] and cross-layer restoration (CLR) [9] to tackle packet layer outages in flexible-grid-based WSONs [10, 11]. However, they did not consider the unique features of P2MP-TRXs. In this work, we study how to recover P2MP-TRX-based WSONs from packet layer failures (*i.e.*, switch outages) cost-effectively by jointly optimizing flow routing and P2MP-TRX configurations. We design three CLR strategies to fully explore the flexibility of P2MP-TRXs, and propose a hybrid hierarchical auxiliary graph and clustering (hHAG-C) algorithm to utilize them for highly-efficient CLR.

II. CLR in P2MP-TRX-based WSONs

We model a P2MP-TRX-based WSON as $G(V, E)$, where V and E are the sets of nodes and fiber links, respectively. In the optical layer, each node $v \in V$ consists of an optical switch, which can realize sub-wavelength switching with the granularity of a frequency slot (FS), and a few P2MP-TRXs. We record each P2MP-TRX s_h in a set $\mathbf{S}$, and define V_h and V_l as the sets of nodes where hub and leaf P2MP-TRXs are deployed, respectively. Here, we have $V_h \cap V_l \neq \emptyset$, because the P2MP-TRXs with certain capacities can serve as either hub or leaf. Hence, a node in the optical layer can be equipped with both hub and leaf P2MP-TRXs. In the packet layer, each node $v \in V$ contains a packet switch. When a switch outage has occurred (as shown in Fig. 1(a)), we find all the affected flows through the failed switch v_o ($v_o \in V$) and record them in a set $\mathbf{R}$[1]. Each flow in $\mathbf{R}$ is denoted as $r_i = (s_i, d_i, x_i)$, where i is its unique index, s_i and d_i are its source and destination, respectively, and x_i denotes its bandwidth in subcarriers (SCs).



(a) Switch outage and CLR

(b) Rerouting over in-service P2MP-TRXs

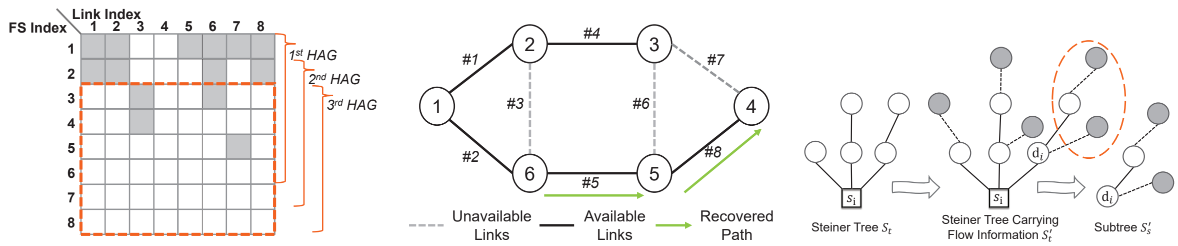
(c) Rerouting over reconfigured P2MP-TRXs

(d) Activating idle P2MP-TRXs

Fig. 1. CLR in a P2MP-TRX-based WSON and proposed CLR strategies.

To evacuate the affected flows cost-effectively, CLR should be applied to rationally use the spare resources in the P2MP-TRX-based WSON, for which we propose three CLR strategies as exampled in Figs. 1(b)-1(d). The first strategy in Fig.

[1]Note that, flows that have v_o as their source or destination nodes cannot be restored, and thus they will not be recorded in $\mathbf{R}$.



(a) Traversing a matrix hierarchically (b) Building auxiliary graph (c) Constructing tree and clustering

Fig. 2. Example of our proposed hHAG-C algorithm.

1(b) tries to utilize the unused SCs on in-service P2MP-TRXs to restore an affected flow and reroute it over fiber links with enough spare FS'. Here, we assume that the outage on *Switch* 2 in Fig. 1(a) interrupts a 75-Gbps flow from *Switch* 6 to *Switch* 4. In Fig. 1(b), the spectrum usage before CLR indicates that the hub P2MP-TRX attached to *Switch* 6 has enough spare SCs to carry the affected flow and the corresponding FS' on *Links* 5 and 8 are available. Hence, we reroute the flow over an in-service P2MP-TRX to restore it.

Fig. 1(c) explains the second strategy, which still tries to use the unused SCs on in-service P2MP-TRXs for CLR but needs to reconfigure their spectrum assignments due to insufficient spare FS' on fiber links. This time, we assume that the affected flow is at 150 Gbps, *i.e.*, the spare FS' on *Links* 5 and 8 become insufficient. Then, the CLR has to reconfigure the hub P2MP-TRX's spectrum assignment as shown in Fig. 1(c). Fig. 1(d) shows the last strategy, which activates idle P2MP-TRXs when in-service ones do not have enough unused SCs to carry an affected flow. The affected flow in Fig. 1(a) is at 400 Gbps, the in-service hub P2MP-TRX on *Switch* 6 cannot support it anymore and we have to activate an idle one there for CLR.

We quantify the strategies' costs as $c_f \cdot n_f + c_l^t \cdot n_l^t$ (first), $c_f \cdot n_f + c_l^t \cdot n_l^t + c_m$ (second), and $c_f \cdot n_f + c_h^t \cdot n_h^t + c_l^t \cdot n_l^t + c_r$ (third), where c_f and n_f are the unit cost and number of newly-activated FS', respectively, c_h^t and n_h^t are the unit cost and number of newly-activated hub P2MP-TRXs of type-t, c_l^t and n_l^t are the unit cost and number of newly-activated leaf P2MP-TRXs of type-t, c_m is the cost of a P2MP-TRX reconfiguration, and c_r is the cost of setting up a new lightpath. Here, the reason why we consider the cost of newly-activated leaf P2MP-TRXs is that since the modulation format used for each SC depends on the length of its lightpath [1], there may be cases in each strategy where the number of SCs used for an affected flow increases after the CLR (*i.e.*, its recovered lightpath is longer than the original one) [12].

III. Algorithm Design

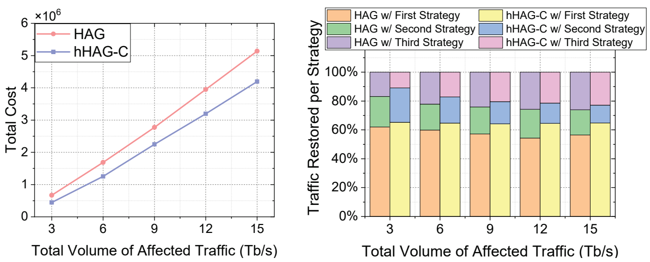
We propose a hybrid hierarchical auxiliary graph and clustering (hHAG-C) algorithm to use the three CLR strategies such that the cost of CLR can be minimized when evacuating all the affected flows in $\mathbf{R}$. Fig. 2 explains the key steps in hHAG-C. We design an approach based on hierarchical auxiliary graph (HAG) to apply the first strategy, while for the second and third strategies, we propose a clustering-based approach to group the affected flows into clusters to find the routing and spectrum assignment (RSA) [10] to restore them. Therefore, hHAG-C includes two sub-procedures.

Step 1 (CLR of Small-sized Flows): We sort the affected flows in $\mathbf{R}$ in ascending order of their bandwidth demands. Then, for each $r_i \in \mathbf{R}$, we traverse all the unused SCs of hub P2MP-TRXs at its source s_i, and build an HAG $G_l(V_l, E_l)$ for it. As shown in Fig. 2(a), we first obtain a matrix to record the FS usages of in-service lightpaths, then traverse the matrix to determine whether a fiber link leaves enough unused FS' for the first strategy, and if yes, we add the link $e \in E$ to $G_l(V_l, E_l)$ (as shown in Fig. 2(b)). After obtain the HAG G_l, we calculate the shortest paths from s_i to d_i in it to store in P. If we have $P \neq \emptyset$, we check the path to get the modulation format for the SCs to use it and determine whether the corresponding in-service hub P2MP-TRX on s_i still has enough SCs to recover r_i with the modulation format. If yes, we just restore r_i accordingly. If r_i cannot be restored in **Step 1**, we put it in set $\mathbf{R}_l$ for being recovered in **Step 2**.

Step 2 (CLR of Large-sized Flows): We divide $\mathbf{R}_l$ into $\{\mathbf{R}_l^{s_i}\}$ based on the source s_i of each flow $r_i(s_i, d_i, x_i) \in \mathbf{R}_l$. For each $\mathbf{R}_l^{s_i} \subseteq \mathbf{R}_l$, we record s_i and d_i of each r_i in it in set S_n, and calculate the Steiner tree S_t in the original topology $G(V, E)$ using S_n as the end node set and s_i as the root (the left subplot of Fig. 2(c)). Then, we modify S_t to a new one S_t' that can represent the information of flows in $\mathbf{R}_l^{s_i}$. As shown in the middle subplot of Fig. 2(c), we add a virtual node (dark solid circle) to connect to each $d_i \in S_n$ to represent the corresponding r_i, while the length of the link (dotted line) between them is set as 0. Then, we check all the hub P2MP-TRXs on s_i (idle P2MP-TRXs can be activated), and cluster the flows in $\mathbf{R}_l^{s_i}$ in S_t' according to the number of unused SCs U_{sc} there. Specifically, we traverse S_t' with the depth-first search, and obtain the subtree S_s' when the total bandwidth demand of flows X_d in S_s' satisfies $\frac{U_{sc}}{2} \leqslant X_d \leqslant U_{sc}$ (the right subplot of Fig. 2(c)). Then, we calculate the Steiner tree S_h in $G(V, E)$ using s_i and destinations in S_s' and the destinations of the original flows of that hub P2MP-TRX as the end node set and get the RSA schemes to restore the flows in S_s' accordingly (by applying the second and third strategies).

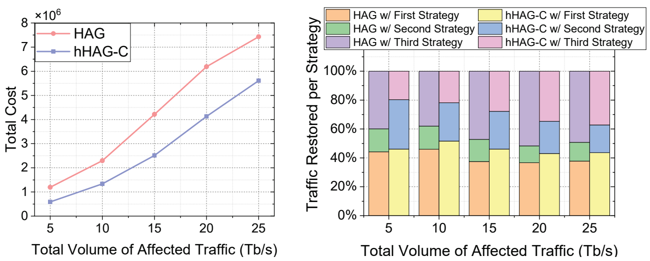
IV. Performance Evaluations

Our simulations use the 24-node US Backbone topology, where each link carries 358 FS', each of which occupies 12.5 GHz. Each SC takes 4 GHz, and thus for a P2MP-TRX at $\{25, 100, 400\}$ Gbps, its maximum spectrum usage is $\{1,2,6\}$ FS', corresponding to $\{1,4,16\}$ SCs. We consider two modulation formats, *i.e.*, DP-16QAM ($\leqslant$ 500 km) and DP-QPSK (others), for each SC to deliver capacities of 25 and 12.5 Gbps, respectively. We simulate light and heavy traffic scenarios, where the demands of flows are within $[1,4]$ and $[5,8]$ SCs, respectively, and in each simulation, we randomly fail a switch to get a set of affected flows. We set $c_f = 1$, $c_m = 1,000$, and $c_r = 100,000$, and the unit cost of a 25/100/400-Gbps P2MP-TRX is $100/200/400$. We use the algorithm without flow clustering (HAG) as the benchmark.



(a) Total cost (b) Traffic restored per strategy

Fig. 3. Results of light traffic scenario.



(a) Total cost (b) Traffic restored per strategy

Fig. 4. Results of heavy traffic scenario.

Fig. 3 shows the simulation results of the light traffic scenario. Fig. 3(a) compares the performance of hHAG-C and HAG in terms of the total cost of CLR, and as expected, the total costs of hHAG-C and HAG increase gradually with the total traffic volume of affected flows. Meanwhile, hHAG-C always provides smaller total costs than HAG, and the gap between their results increases with the total volume of affected flows. Specifically, hHAC-C achieves a maximum cost saving of 33% and an average cost saving of 23% over HAG. To further explore the factors that can affect the algorithms' performance, Fig. 3(b) shows the distributions of the CLR strategies used by the algorithms. We can see that hHAG-C restores more flows with the first and second strategies than HAG, which reveals that efficiently utilizing the spare resources on the P2MP-TRXs and fiber links helps to provide more economical CLR schemes.

Fig. 4 shows the results of the heavy traffic scenario, and similar trends can be observed as those in Fig. 3. In Fig. 4(a), hHAC-C achieves a maximum cost saving of 51% and an average cost saving of 38% over HAG. The reason for this phenomenon can still be found in Fig. 4(b). With hHAG-C, an average of 72% of the affected flows are restored by the first and second strategies, compared to only 55% by HAG. Note that, the higher bandwidth demands in the heavy traffic scenario make a larger portion of the affected flows unable to be restored by the first strategy. As a result, there is a greater reliance on the second and third strategies. hHAG-C effectively aggregates the affected flows to find cost-efficient CLR strategies for them, thereby avoiding unnecessary reconfiguration and activation of P2MP-TRXs to save the cost of CLR.

Table I lists the running time of the algorithms and indicates that hHAG-C can obtain CLR schemes for all the affected flows within 2.7 seconds, while in comparison, HAG can take ~21.8 seconds at most. This is attributed to the sub-procedure of flow clustering in hHAG-C, which enables restoring multiple flows simultaneously, while HAG can only recover one flow each time, resulting in a slower execution speed.

TABLE I
RUNNING TIME OF ALGORITHMS (SECONDS)

Light Traffic Scenario			*Heavy Traffic Scenario*		
Traffic in $\mathbf{R}$ (Tb/s)	hHAG-C	HAG	Traffic in $\mathbf{R}$ (Tb/s)	hHAG-C	HAG
3	0.8842	2.4355	5	0.8413	3.0279
6	1.599	5.1031	10	1.3815	5.6377
9	1.6523	7.9543	15	1.9522	10.2244
12	2.283	12.3026	20	2.1764	15.7474
15	2.611	16.3249	25	2.6311	21.7667

V. CONCLUSION

We designed three CLR strategies to fully explore the flexibility of P2MP-TRXs for recovering P2MP-TRX-based WSONs from packet layer failures, and proposed an hHAG-C algorithm to use the strategies for highly-efficient CLR. Extensive simulations confirmed the effectiveness of our proposal.

REFERENCES

[1] D. Welch *et al.*, "Point-to-multipoint optical networks using coherent digital subcarriers," *J. Lightw. Technol.*, vol. 39, pp. 5232–5247, Aug. 2021.

[2] 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.

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

[4] Z. Zhu *et al.*, "Impairment- and splitting-aware cloud-ready multicast provisioning in elastic optical networks," *IEEE/ACM Trans. Netw.*, vol. 25, pp. 1220–1234, Apr. 2017.

[5] R. Li, Q. Lv, and Z. Zhu, "On the network planning of wavelength switched optical networks with P2MP transceivers," *J. Lightw. Technol.*, vol. 42, pp. 24–36, Jan. 2024.

[6] 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.

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

[8] 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.

[9] 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.

[10] 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.

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

[12] Z. Zhu *et al.*, "Energy-efficient translucent optical transport networks with mixed regenerator placement," *J. Lightw. Technol.*, vol. 30, pp. 3147–3156, Oct. 2012.