Integrating FlexE with Wavelength-Selective Optical Networks: P2P *versus* P2MP Transceivers

(Invited Paper)

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

***Abstract*—Traditional Flexible Ethernet (FlexE) architecture based on point-to-point transceivers (P2P-TRXs) struggles to efficiently accommodate the prevailing hub-and-spoke (H&S) traffic in today's metro/access networks due to their limited traffic aggregation capability. Recently, coherent point-to-multipoint transceivers (P2MP-TRXs), with their capability to simultaneously reach multiple destinations and their flexible subcarrier-level scheduling, have exhibited inherent advantages for handling H&S traffic. Moreover, P2MP-TRXs offer comparable unit costs to P2P-TRXs, making them a promising alternative. In this paper, we study the synergistic benefits between FlexE and P2MP-TRXs, namely, FlexE-P2MP, and compare it with two baseline architectures: FlexE with P2P-TRXs (FlexE-P2P) and link aggregation groups with P2MP-TRXs (LAG-P2MP). Extensive simulations validate the advantages of FlexE-P2MP in terms of resource efficiency, cost-effectiveness, and scalability.**

***Index Terms*—Flexible Ethernet (FlexE), Point-to-multipoint transceivers (P2MP-TRX), Architecture comparison.**

I. Introduction

Nowadays, IP over optical transport networks (IPoOTNs) face significant challenges in optimizing bandwidth allocation and adapting to changing traffic patterns [1–9]. Flexible Ethernet (FlexE) effectively addresses the mismatch between variable media access control (MAC) rates and fixed physical (PHY) capacities in traditional Ethernet by introducing a FlexE shim between MAC and PHY layers to support bonding, subrating, and channelization. It uses time-division multiplexing (TDM) to divide PHY capacity into fine-grained transmission slots (TS'), enabling precise resource allocation, improving bandwidth utilization, and ensuring deterministic delay [10].

Concurrently, applications such as cloud computing and distributed learning are driving network traffic toward hub-and-spoke (H&S) patterns [11, 12]. However, existing FlexE paradigms still primarily rely on point-to-point transceivers (P2P-TRXs) [13–16], which struggle to support H&S traffic efficiently. Fortunately, recently emerging coherent point-to-multipoint transceivers (P2MP-TRXs) employ digital subcarrier multiplexing (DSCM) to distribute high-speed traffic from the hub to multiple leaf nodes via low-rate Nyquist subcarriers (SCs), significantly reducing transceiver requirements while simplifying network control and management [17]. As shown in Fig. 1(b), serving multiple destinations might require activating additional P2P-TRXs even when the TS' in a FlexE group remain under-utilized. This single-destination constraint significantly limits FlexE's scheduling flexibility and resource efficiency. The integration of FlexE with P2MP-TRXs (FlexE-P2MP) combines FlexE's fine-grained TS' scheduling with P2MP-TRXs' multi-destination capabilities. As shown in Fig. 1(c), this approach enables the allocation of traffic within a FlexE group to multiple destinations through a single P2MP-TRX, substantially improving overall resource utilization.

This paper examines the synergistic benefits of FlexE-P2MP through comparative analysis against two benchmark architectures: FlexE with P2P-TRXs (FlexE-P2P) and link aggregation groups with P2MP-TRXs (LAG-P2MP). Our comprehensive simulations across various traffic scenarios demonstrate FlexE-P2MP's superior performance in terms of resource utilization, adaptability, and cost-effectiveness.

II. Operation Principle

A. IPoOTN Model

We model an IPoOTN as a graph $G(V, E)$, where V and E represent the sets of nodes and links, respectively. Each client stream in the IPoOTN is denoted by $c_i(s_i, d_i, x_i)$, where i is the unique identifier of the stream, s_i, d_i, and x_i denote the source, destination, and required bandwidth, respectively. Both Ethernet trunk (Eth-Trunk) and FlexE group bundle four 100 Gbps PHYs into a single logical link of equivalent capacity to support higher-rate client streams [13–16].

For FlexE, we adopt the FlexE-terminal-based approach to achieve efficient spectrum utilization. In this architecture, a FlexE shim is inserted between MAC and PHY layers to divide each 100 Gbps PHY into 20 TS' of 5 Gbps each. Client streams are mapped onto a set of TS', which are then identified and switched at a transport box (T-Box) by another FlexE shim before being further mapped into OTN for transmission [15].

In addition, P2P-TRXs adopt bandwidth-variable transponders (BV-Ts) for flexible and dynamic bandwidth configuration [18, 19]. Different types of P2MP-TRXs can activate 1/4/16 SCs, achieving data rates of 25/100/400 Gbps under DP-16QAM modulation format [17]. During the provisioning of client streams, we must simultaneously determine a feasible path, an appropriate set of TS', and a contiguous spectrum segment. The objective is efficient resource distribution to satisfy all transmission requirements while minimizing the total expenditure on transceivers, spectrum and so forth.

Our analysis compared three alternative architectures:

1) LAG-P2MP is the P2MP-based wavelength-selective optical network (WSON) without FlexE integration, where LAG distributes client streams across PHYs of the Eth-Trunk using load balancing, then transmits to respective destinations through P2MP-TRXs.

2) FlexE-P2P refers to the WSON architecture that integrates FlexE with P2P-TRXs, where each P2P-TRX can only transmit data to a single destination.
3) FlexE-P2MP leverages both the fine-grained TS' allocation of FlexE and the multi-destination transmission capability of P2MP-TRXs to efficiently serve client streams.

B. Architecture Comparison

To explain the difference among the three architectures clearly, we examine their performance across two traffic scenarios. *Scenario* 1 involves five client streams: $c_1(A,B,10)$, $c_2(A,B,40)$, $c_3(A,B,75)$, $c_4(A,C,75)$, and $c_5(A,D,125)$ (data rates in Gbps). *Scenario* 2 introduces traffic changes by removing c_5 and adding two new streams: $c_6(A,D,100)$ and $c_7(A,E,100)$. For clarity in this analysis, we adopt the same routing scheme and ignore their effects on modulation formats.

In *Scenario* 1, as shown in Fig. 1(a), LAG-P2MP distributes the client streams onto an Eth-Trunk and then assigns $\{1,2,3,3,5\}$ SCs to the five streams through a P2MP-TRX, respectively. Notably, c_1 and c_2 under-utilize their assigned SC capacities (*i.e.*, using only 40% and 80% of them, respectively), which reduces the overall SC utilization to 84%. For FlexE-P2P (Fig. 1(b)), the five streams are respectively allocated $\{2,8,15,15,25\}$ TS' within the FlexE group and transmitted to separate destinations via three P2P-TRXs. Due to the single-destination constraint of P2P-TRXs, the PHY utilization of the FlexE group is significantly reduced, and the number of required TRXs and T-Boxes increases, leading to additional capital expenditure (CAPEX). For FlexE-P2MP (Fig. 1(c)), we still allocate $\{2,8,15,15,25\}$ TS' for five streams, respectively, but with the multi-destination capability of P2MP-TRXs, all the streams can reach their destinations through a single P2MP-TRX. This significantly reduces the number of TRXs and T-Boxes. Moreover, compared to LAG-P2MP, FlexE enables time-sharing on spectrum resources: c_1 and c_2 alternate transmit over *SC* 1 using 2 and 3 TS', while the remaining 5 TS' of c_2 are transmitted over *SC* 2, fully utilizing the available SC capacity.

When traffic changes in *Scenario* 2, as illustrated in Fig. 2(a), the limited SC utilization in LAG-P2MP prevents the activated P2MP-TRX from accommodating both c_6 and c_7. After assigning 4 SCs to c_6, a new P2MP-TRX has to be deployed for c_7. Similarly, FlexE-P2P (Fig. 2(b)) requires the activation of an additional P2P-TRX to support c_7, since it targets a new destination (*Node* E). In contrast, FlexE-P2MP (Fig. 2(c)) accommodates both c_6 and c_7 without new hardware deployment, due to the fine-grained TS' scheduling offered by FlexE and the multi-destination capability of P2MP-TRXs. This demonstrates the superior adaptability of FlexE-P2MP in handling dynamic traffic variations.

Overall, we summarize the resource usages across the two scenarios as follows. In *Scenario* 1, LAG-P2MP uses one TRX and one T-Box, and activates 14 SCs, FlexE-P2P uses 3 TRXs and 2 T-Boxes, and FlexE-P2MP uses one TRX and one T-Box, and activates 13 SCs. In *Scenario* 2, LAG-P2MP uses 2 TRXs and one T-Box, and activates 17 SCs, FlexE-P2P uses 4 TRXs and 2 T-Boxes, and FlexE-P2MP uses one TRX and one T-Box, and activates 16 SCs. These results indicate that introducing P2MP-TRXs improves PHY utilization and reduces the TRXs and T-Boxes required. Furthermore, FlexE enhances SC utilization in P2MP-TRXs, lowering the SCs needed. Hence, FlexE-P2MP achieves the highest resource efficiency, providing a more cost-effective solution for IPoOTN.

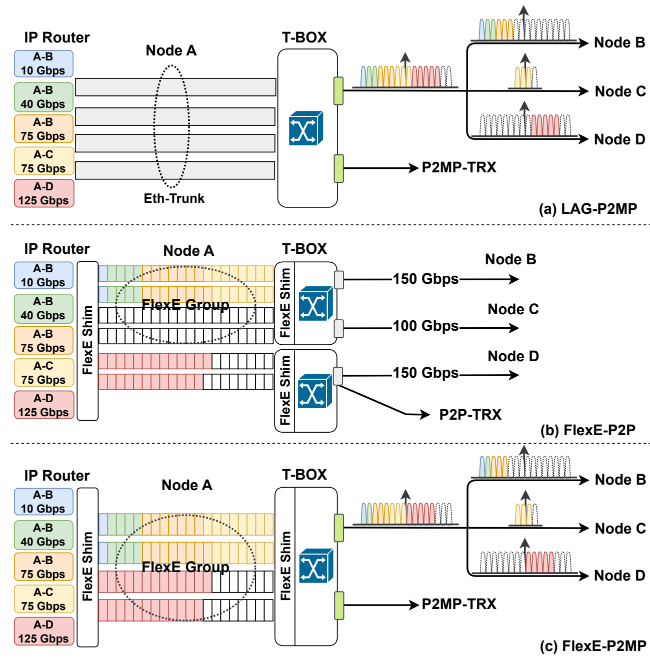


Fig. 1. *Scenario* 1 for traffic distribution.

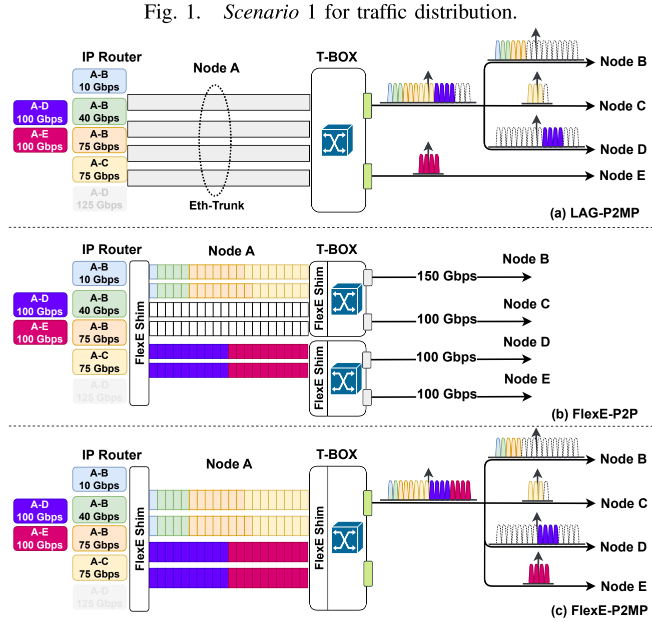
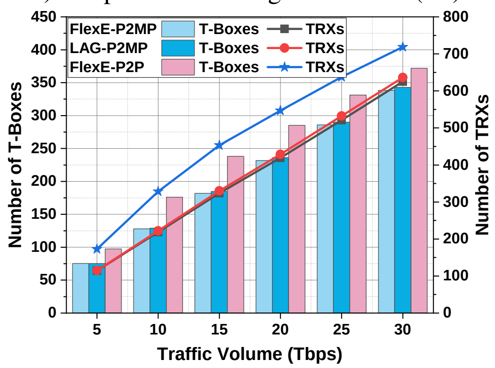


Fig. 2. *Scenario* 2 for traffic distribution.

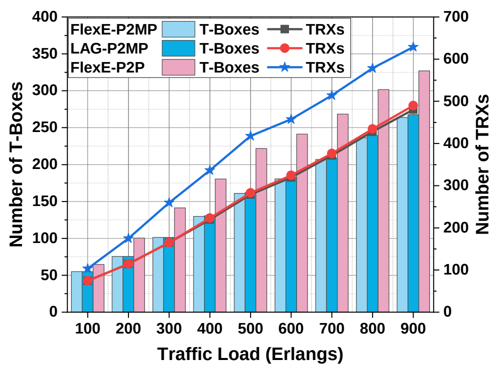
III. Performance Evaluations

We evaluate the performance of LAG-P2MP, FlexE-P2P, and FlexE-P2MP under different traffic volumes on the 24-node US backbone network topology [20]. Each Eth-Trunk and FlexE group is assumed to bundle four 100 Gbps PHYs with a T-Box. Each T-Box is equipped with two TRXs [16].

For P2MP-TRXs, different types are considered, capable of activating 1/4/16 SCs. Each SC occupies 4 GHz and can transmit at 25 Gbps or 12.5 Gbps per SC depending on the modulation format: DP-16QAM (for transmission distances $\leq$ 500 km) or DP-QPSK [21–23]. Test case 1 is for static traffic, and the source and destination of each client stream are randomly selected from V. The data rates are uniformly drawn from $\{10, 40, 25 \cdot n\}$ Gbps, where we set $n \in [1, 8]$. Test case 2 generates client streams dynamically, with their arrival time and life-time following the Poisson traffic model. To ensure fairness in the architectural comparison, all the routing paths and P2MP trees are computed based on the shortest-path mechanism, and resource allocations (including TS' and SC assignments) are performed using the first-fit (FF) strategy.



(a) Test case 1



(b) Test case 2

Fig. 3. Performance comparison of three architectures.

Fig. 3(a) compares the TRXs and T-Boxes used by the three architectures in test case 1. Consistent with our previous discussions, FlexE-P2MP utilizes the fewest TRXs and T-Boxes, followed by LAG-P2MP, while FlexE-P2P requires the most. It's important to note that for LAG-P2MP, we assume full bandwidth utilization of Eth-Trunks, but in actual networks, the uneven hash distribution in LAG load balancing typically limits bandwidth utilization to 70-80% [10], meaning real-world performance would be even worse. Fig. 3(b) exhibits the same trend as that of Fig. 3(a), highlighting that FlexE-P2MP is also well-adapted to dynamic traffic conditions.

IV. Conclusion

In this work, we conducted a comprehensive comparison between FlexE-P2MP and two benchmark architectures, LAG-P2MP and FlexE-P2P, under static and dynamic traffic scenarios. Extensive simulations demonstrate that FlexE-P2MP consistently outperforms the benchmarks in terms of resource utilization, cost-effectiveness, and adaptability, making it a promising and scalable solution.

Acknowledgments

This work was supported by the NSFC project 62371432.

References

[1] 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.

[2] Y. Yin *et al.*, "Spectral and spatial 2D fragmentation-aware routing and spectrum assignment algorithms in elastic optical networks," *J. Opt. Commun. Netw.*, vol. 5, pp. A100–A106, Oct. 2013.

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

[4] C. Chen *et al.*, "Demonstrations of efficient online spectrum defragmentation in software-defined elastic optical networks," *J. Lightw. Technol.*, vol. 32, pp. 4701–4711, Dec. 2014.

[5] 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.

[6] W. Lu, Z. Zhu, and B. Mukherjee, "On hybrid IR and AR service provisioning in elastic optical networks," *J. Lightw. Technol.*, vol. 33, pp. 4659–4669, Nov. 2015.

[7] P. Lu and Z. Zhu, "Data-oriented task scheduling in fixed- and flexible-grid multilayer inter-DC optical networks: A comparison study," *J. Lightw. Technol.*, vol. 35, pp. 5335–5346, Dec. 2017.

[8] 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.

[9] X. Dong *et al.*, "On scheduling DML jobs in all-optical DCNs with in-network computing," in *Proc. of GLOBECOM 2024*, pp. 1–6, Dec. 2024.

[10] FlexE White Paper. [Online]. Available: https://www.oiforum.com/wp-content/uploads/2019/01/OIF_FlexE_White_Paper.pdf

[11] 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.

[12] K. Wu, P. Lu, and Z. Zhu, "Distributed online scheduling and routing of multicast-oriented tasks for profit-driven cloud computing," *IEEE Commun. Lett.*, vol. 20, pp. 684–687, Apr. 2016.

[13] 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.

[14] 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. Manage.*, vol. 18, pp. 3691–3705, Sep. 2021.

[15] 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.

[16] M. Wu, N. L. S. da Fonseca, and Z. Zhu, "Dynamic cross-layer restoration to resolve packet layer outages in flexe-over-eons," *IEEE Trans. Netw. Serv. Manage.*, vol. 19, pp. 2600–2611, Sep. 2022.

[17] D. Welch *et al.*, "Digital subcarrier multiplexing: Enabling software-configurable optical networks," *J. Lightw. Technol.*, vol. 41, pp. 1175–1191, Feb. 2023.

[18] N. Sambo *et al.*, "Next generation sliceable bandwidth variable transponders," *IEEE Commun. Mag.*, vol. 53, pp. 163–171, Feb. 2015.

[19] X. Chen *et al.*, "DeepRMSA: A deep reinforcement learning framework for routing, modulation and spectrum assignment in elastic optical networks," *J. Lightw. Technol.*, vol. 37, pp. 4155–4163, Aug. 2019.

[20] M. Wu *et al.*, "On the cross-layer restoration to address packet layer failures in P2MP-TRX-based WSONs," *J. Opt. Commun. Netw.*, vol. 17, pp. 95–109, Feb. 2025.

[21] Y. Zhang *et al.*, "Planning of survivable wavelength-switched optical networks based on P2MP transceivers," *IEEE Trans. Netw. Serv. Manag.*, vol. 20, pp. 2331–2342, Sept. 2023.

[22] Q. Lv, R. Li, and Z. Zhu, "On the survivable multilayer planning of filterless optical networks with P2MP transceivers," *J. Lightw. Technol.*, vol. 42, pp. 1786–1797, Mar. 2024.

[23] 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.