

Optimizing Light-Tree and Edge Resource Placement for Diverse Latency Needs

P. Soumplis^{1,2}, K. Yiannopoulos^{1,3}, K. Christodouloupoulos^{1,4}, E. Varvarigos^{1,2}

¹*Institute of Communication and Computer Systems, Zografou, Greece*

²*School of Electrical and Computer Engineering, National Technical University of Athens, Zografou, Greece*

³*Department of Informatics and Telecommunications, University of the Peloponnese, Tripoli, Greece*

⁴*Department of Informatics and Telecommunications, National and Kapodistrian University of Athens, Greece*
e-mail: soumplis@mail.ntua.gr

ABSTRACT

Coherent Point-to-MultiPoint (P2MP) transceivers are gaining attention for their efficacy in aggregating diverse workloads through optical light-trees, particularly in extended access scenarios where P2MP hub nodes integrate edge computing capabilities. This work addresses the challenge of optimizing the placement of both computing and networking resources to serve various workload quality classes, which can be suboptimal when treated independently. This is done taking into account the latency and network capacity requirements of the different classes, the optical connections physical layer performance, as well as computational costs.

Keywords: P2MP connectivity, light trees, edge computing.

1. INTRODUCTION

In the recent years, we have witnessed a transformation in the workload requirements, driven by the emergence of new applications with stringent computing and networking needs. This surge encompasses not only mobile and IoT services but also extends to autonomous driving, virtual and augmented reality, and the advancements within Industry 4.0. Each of these domains includes applications that require real-time processing and ultra-low latency, presenting challenges for traditional cloud computing and networking models [1].

To tackle these obstacles, a significant shift towards edge computing has emerged as a promising solution. Edge computing manages to significantly reduce latency and enhance reliability by decentralizing processing and placing computational resources closer to the end-users and the data sources, at the network's periphery. A variety of edge platforms have been proposed, from small-scale cloudlets, to FPGAs, ASICs, and GPUs. Each one is positioned at different layers to meet the different computational capacities, costs, and energy needs.

Typically, lower capacity devices are deployed on-premise or at near edge to handle tasks requiring low latency and small computing power. Meanwhile, computation-strong resources are positioned at the far edge, located at telecom operators' Points Of Presence (POP), such as central offices in mobile networks or ISP nodes, facilitating the execution of larger-scale applications. Each POP location is associated with distinct expenses including rent, setup and operational costs, which are critical factors when performing network planning and budgeting.

Furthermore, network architecture and devices have evolved to support this increased demand for low-latency and high-capacity data transmission. The introduction of optical coherent Point-to-Multipoint (P2MP) technologies emerged as an alternative to traditional Point-to-Point (P2P) configurations, offering advantages such as lower costs, enhanced flexibility, and improved efficiency [2]. Coherent P2MP transceivers use Digital Sub-Carriers (DSCs), where each sub-carrier can be independently allocated, enabling thus a single high-capacity transceiver to simultaneously transmit/receive traffic from multiple lower-rate sources forming a light-tree. This suits the typical "hub and spoke" architecture that is used in metro-aggregation horseshoe or ring networks, where a single high-capacity "hub" node communicates with several "spoke" nodes, while spokes do not directly communicate, since inter-spoke traffic is low and does not justify the use of additional transceivers.

In the reference architecture, a leaf (also referred to as aggregation) node aggregates the traffic from various access points, like mobile sites or fixed access connections (e.g., several OLTs at a leaf node). Then, these nodes are connected in a horseshoe to two metro-core nodes at each horseshoe endpoint (or to the same metro-core node in a ring). This architecture suits the deployment of P2MP transceivers; the metro-core nodes at horseshoe endpoints become the hubs with P2MP transceivers, and the aggregation nodes become the spokes that deploy multiple small-rate transceivers to communicate only with a hub, forming light-trees over the horseshoe. Hence, the externally destined traffic from the aggregation nodes is terminated at the metro-core node at the horseshoe end, and is electrically aggregated and forwarded through P2P connections to one of the backbone POPs.

However, the use of coherent P2MP transceivers supports transmission over greater distances, enabling the disaggregation of the reference horseshoe architecture and the movement of the hub/P2MP nodes inside the metro network as opposed to each horseshoe's end-points. This allows for more flexible light-tree operations and increases the gains from multiplexing different traffic sources. As edge resources can be located at various levels

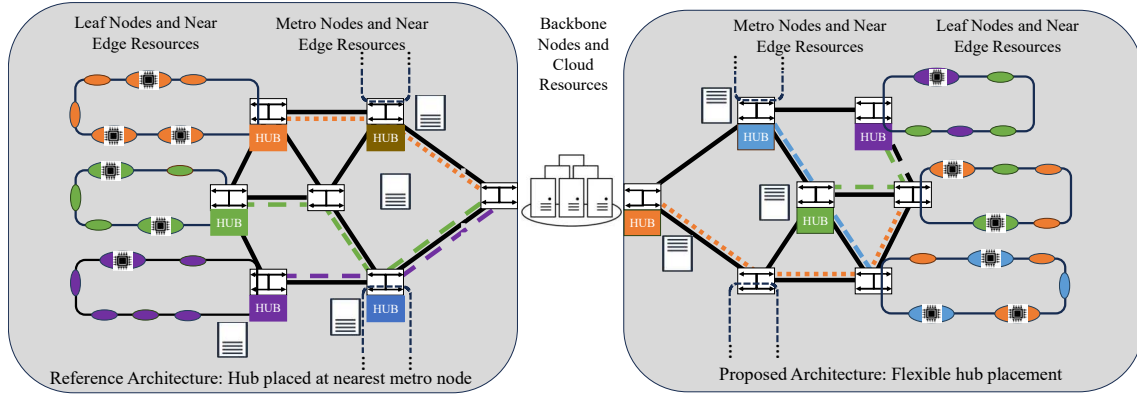


Figure 1. The reference (left side) and proposed (right side) architectures for light-tree communication.

within the metro part of the network, the placement of P2MP transceivers should, in addition to physical layer feasibility (OSNR) and capacity of the placed transceivers, take into consideration the availability of different metro nodes and their related costs, but also the latency requirements of the different classes of workloads. When P2MP transceivers are placed wisely along with computing resources, not only is the latency of the provided services decreased and the multiplexing gains increased, but also larger traffic volumes can be supported with the different workloads terminating at the different layers of the infrastructure.

In this work, we examine the problem of joint placement of P2MP transceivers and the activation of computing nodes, taking into consideration the latency requirements of the various workloads. The paper is structured as follows: Section II details the system model, Section III discusses the latency-aware greedy mechanism for transceiver and edge location selection, Section IV presents simulation results, and Section V the conclusions.

2. SYSTEM MODEL

The system model under study, illustrated in Figure 1, presents two distinct architectures: the reference architecture on the left side and the proposed architecture on the right side. In both architectures, at the lowest level, leaf nodes aggregate the traffic from multiple access sites, including fixed OLT PONs and mobile sites. These act as the initial aggregation points, which either forward the traffic to higher hierarchy levels or terminate it locally based on the traffic's requirements (ultra-low latency) and availability of computing resources.

Following the leaf nodes in the network hierarchy are the hub nodes, which function as the second level of aggregation. Their role is twofold: either to route the traffic to local computing resources for immediate processing, or forward it to the core network. Hence, they enhance the network efficiency by optimizing the flow of data and reducing latency, ensuring that the routing of traffic is performed with respect to its latency requirements. A key optimization parameter is the location of the hubs: in the reference architecture they are always located at horseshoe endpoints which connect to the metro-network (left side in Figure 1), while in the proposed architecture these can be moved internally to the metro. In the following we focus our discussion on the proposed architecture.

The communication of hub nodes with leaf nodes is performed with P2MP transceivers by establishing light-trees so that the hub communicates with multiple leaf nodes. The transmission between the hub and a leaf node is performed over a path that introduces noise and impairments, influencing the number of sub-carriers required to achieve a feasible transmission at the desired rate. Hence, the capacity of the transceiver at the hub, characterized by specific parameters such as total capacity and number of sub-carriers, along with predefined feasibility thresholds, determine, which leaf nodes the hub can serve.

To enable the communication through light-trees P2MP, transceivers that utilize digital sub-carriers (DSCs) are required. Since DSCs can be independently allocated, a single P2MP transceiver can concurrently serve traffic from multiple leaf nodes. These P2MP transceivers are characterized by their acquisition cost, baud rate, and supported modulation formats, such as QPSK. Consequently, the feasible maximum transmission distance is determined by the OSNR threshold, which varies according to the specific modulation format employed.

In the proposed architecture (right side in Figure 1) in the “downstream” direction of the light-tree, the transceivers placed at the hub, broadcast their SCs to all connected horseshoes via modified ROADMs [3] and the leaves perform demultiplexing to isolate their data. In the “upstream”, the SCs from all leaves are passively combined at the horseshoes through the modified ROADMs to reach the hub node. Additionally, P2P transponders can forward the traffic further in the metro network due to insufficient computing or networking resources.

The computing resources deployed at the edge, vary in load and the related execution cost. Lower activation costs are observed for far-edge resources (regional and metro) due to the ability to leverage pre-existing infrastructures that may already be partially configured for other purposes. In contrast, activation expenditures for

near-edge nodes (leaf nodes, lower-level metro-core nodes) are higher. These costs are related to the investments required to build or rent the necessary infrastructure. Additionally, the budget for these nodes also includes energy constraints influenced by factors such as the local energy policy.

We consider that we know the traffic demands starting at the access nodes belonging to traffic classes with different levels of latency and capacity requirements. The demands in both uplink and downlink are characterized by the size of the demand (in Gb/s), assumed to be equal for both directions, and a service latency that also denotes the priority of the execution, distinguished mainly into three classes (i) low latency/ high priority, (ii) medium latency and also priority and (iii) high latency / low priority.

3. LATENCY AWARE P2MP TRANCEIVER AND EDGE LOCATION SELECTION

In this section, we present the proposed latency-aware computing and networking resources placement mechanism. This mechanism's inputs include information regarding (i) the network topology and related physical layer parameters for the access, aggregation (horseshoe) and metro segments; (ii) detailed specifications of P2MP and P2P optical transceivers, including their supported modulation formats, baud rates, and Optical Signal-to-Noise Ratio (OSNR) thresholds; (iii) economic parameters associated with activating computing resources at each node, such as installation and operational costs; and (iv) a set of demands with their data volume and latency class.

Initially, the algorithm calculates for each demand the shortest path towards all the metro nodes and sorts the demands based on their latency class to prioritize those requiring the lowest latency. Then, for each demand and based on the paths and physical layer network parameters, it calculates the path feasibility in terms of physical performance/OSNR, also considering the demand's latency requirement (i.e. propagation latency calculated according to the path distance). Based on those calculations, it identifies the set of nodes it can reach with P2MP transceivers. Note that the high latency types of traffic are assumed to terminate outside of the metro segment and P2P transceivers should be deployed at the hub location to forward that traffic further to core/backbone nodes.

After calculating the potential nodes for placing the P2MP transceivers (and subsequently P2P if needed), the best nodes, those that incur the lowest computing costs, are identified in a way similar to the weighted set cover problem that aims at covering all demands by activating the fewest possible nodes while simultaneously minimizing the incurred costs. This combinatorial problem is inherently complex, as it belongs to the class of NP-complete problems [4], and is addressed using a sub-optimal greedy algorithm.

In particular, this is done in an iterative manner, where in each iteration, the greedy algorithm places a transceiver at the node that covers the maximum number of uncovered demands and assigns the demands of the leaf nodes to the transceiver placed at the selected hub node, a process that is repeated until covering all demands. The output of the algorithm consists of the placement of the transceivers (P2MP or P2P) at the different nodes, together with the configuration of the computing and networking resources, detailed in terms of which nodes are connected, the types of connections (P2MP or P2P) and the connections that utilize computing resources within a node.

4. SIMULATION RESULTS

We performed a number of simulations to evaluate the performance of the proposed architecture and the latency-aware P2MP transceiver and edge-cloud location selection mechanism and compare it to that of the reference architecture. We used Telefonica's Spanish network topology (domain A), featuring six backbone nodes. Each metro-core node in this topology was connected to a horseshoe with 10 leaf nodes, generating SC requests within [0,20], distributed as 10% low latency (0.5 usec), 40% medium latency (2 msec), and 50% high latency (10 msec). We assumed that P2MP transceiver costs double as capacities quadruple [2], and ranged them from 1 to 3 arbitrary units (a.u.), while P2P transceiver costs were set at 90% of P2MP costs [5]. We assumed a 3 dB OSNR degradation per 80 km, in line with the GN model's predictions [6] for nonlinear effects due to close SC spacing (~4 GHz). The activation costs for network resources were set at 1.96 a.u. for near edge, 1.44 a.u. for far edge, and 1 a.u. for backbone nodes, with operational costs proportional to the transmission rate across different levels.

An OSNR budget value of 0 dB corresponds to the baseline scenario and it requires the placement of a hub at each horseshoe end/ the nearest metro-core node and serves the leaf nodes through P2MP transceivers. As OSNR budget values increase, they allow greater transmission reaches, which results in placing the hub nodes higher in the hierarchy. This enables workloads to be served through fewer P2P transceivers (Figure 2.a), effectively reducing activation and operational nodes' costs (Figure 2.c) while improving network efficiency. Figure 2.b highlights the contribution of P2MP transceivers in selecting fewer nodes to place computing resources, justified by their ability to achieve higher multiplexing gains despite their higher costs. As OSNR budgets increase, it becomes feasible to deploy P2MP root transceivers deeper within the metro-core network, significantly reducing the reliance on P2P connections. Finally, Figure 2.d presents the variance in the number of light-trees based on the type of costs considered: networking or computing.

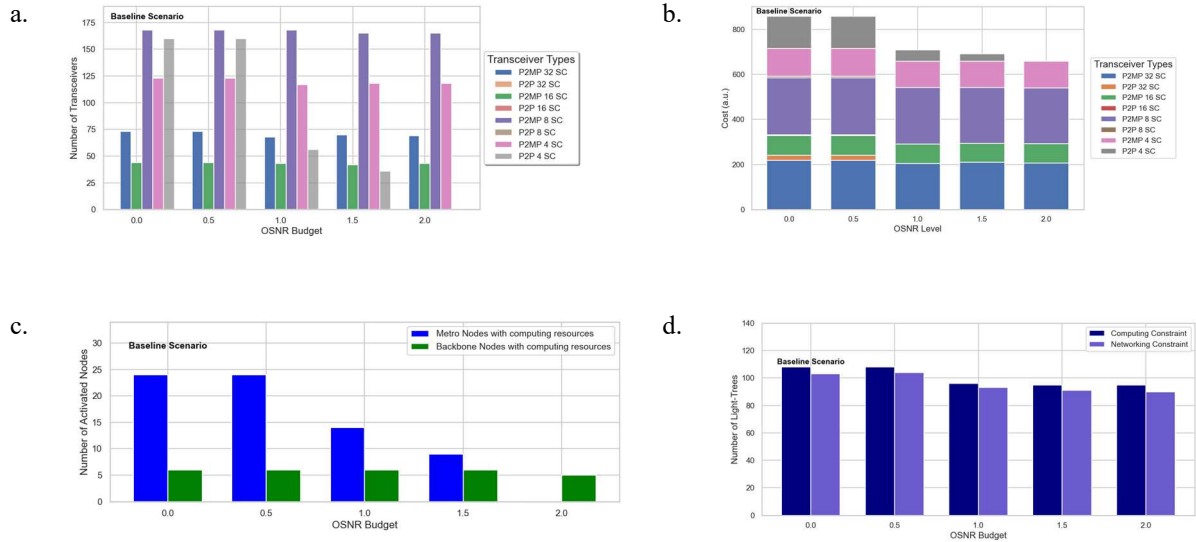


Figure 2. The impact of varying OSNR budgets on (a) the number and type of transceivers required, (b) their associated costs, (c) the number of activated metro and backbone nodes, and (d) the number of light-trees.

When computing costs are prioritized, demands consolidate on fewer nodes, with an increased number of light-trees, demonstrating the suitability of P2MP light-trees in supporting edge computing models and showcasing the benefits of the latency aware computing and networking device placement mechanism in supporting the flexible light-tree allocation.

5. CONCLUSIONS

This paper proposes a computing resource aware approach for optimizing the placement of P2MP transceivers in edge networks to meet diverse latency requirements efficiently. By jointly considering the placement of these resources and the latency needs of various workloads, our approach aims to enhance service delivery and maximize resource utilization. Through simulations on Telefonica's Spanish network topologies, we demonstrated the effectiveness of our latency-aware strategy compared to a static baseline scenario. These findings underscore the significance of integrating edge computing and networking technologies to address the evolving demands of modern applications effectively.

ACKNOWLEDGEMENTS

This research has been funded by the EU H2022 ALLEGRO project (Grant agreement ID: 101092766) and was also partially implemented in the framework of HFRI under the National Recovery and Resilience Plan "Greece 2.0" funded by the European Union – NextGenerationEU (HFRI Project Number: 16828)

REFERENCES

- [1] V. Farhadi et al., "Service Placement and Request Scheduling for Data-Intensive Applications in Edge Clouds," in *IEEE/ACM Transactions on Networking*, vol. 29, no. 2, pp. 779-792, April 2021.
- [2] J. Bäck, A. Napoli, E. Riccardi, et al., "A Filterless Design with Point-to-multipoint Transponders for Cost-Effective and Challenging Metro/Regional Aggregation Topologies," *ONDM 2022*, pp. 1-6.
- [3] K. Yiannopoulos, et al., "Enabling Extended Access with Light-Trees and Point-To-Multipoint Coherent Transceivers," in *Proc. Optical Network Design and Modelling (ONDM)*, 2024, to appear.
- [4] J. Guo and R. Niedermeier, "Exact algorithms and applications for Tree-like Weighted Set Cover," *Journal of Discrete Algorithms*, vol. 4, no. 4, pp. 608-622, Dec. 2006.
- [5] N. Skorin-Kapov, F. J. Moreno Muro, et al., "Point-to-Multipoint Coherent Optics for Re-thinking the Optical Transport: Case Study in 5G Optical Metro Networks," *ONDM 2021*, pp. 1-4.
- [6] P. Poggiolini et al., "Analytical and Experimental Results on System Maximum Reach Increase Through Symbol Rate Optimization," in *Journal of Lightwave Technology*, vol. 34, no. 8, pp. 1872-1885, 15 April 2016.