Optimized Distributed Unit (DU) Placement under Physical Constraints for 6G x-Haul

(Article begins on next page)

Optimized Distributed Unit (DU) Placement under Physical Constraints for 6G x-Haul

Ahtisham Ali*†, Sanwal Zeb†, Muhammad Umar Masood†, Andrea Rosso†, Gulmina Malik†,
Michela Pollone*, Alessandro Galardini*, Vittorio Curri†
*Consorzio TOP-IX, Torino, Italy †DET, Politecnico di Torino, Torino, Italy
Email: ahtisham.ali@top-ix.org

***Abstract*—The disaggregation of radio access networks (RAN) into radio units (RUs), distributed units (DUs), and central units (CUs) enables flexible and virtualized deployments, but imposes stringent constraints on the underlying transport network. In converged metro-access optical networks, DU placement must be carefully designed to ensure feasible RU–DU–CU connectivity under joint strict latency and quality-of-transmission (QoT) constraints. In this work, we investigate DU placement as an optimization problem aimed at maximizing feasible RU–CU routes. A GNPy-based optical simulation framework is used to evaluate the feasibility of RU–DU–CU paths under coherent DP-16QAM transmission and functional split 7.2 latency constraints while comparing exhaustive, greedy, and genetic algorithm (GA)-based strategies to quantify the trade-off between feasible routes optimization and computational complexity. Results show that full RU–CU coverage can be achieved with only two optimally placed DUs using GA and Greedy-Coverage in our considered scenario. Among the evaluated approaches, a coverage-driven greedy approach achieves optimal performance with much lower complexity than genetic and exhaustive approach. These findings show that a network-aware DU placement heuristic provides an efficient and scalable solution for practical disaggregated RAN deployments.**

***Index Terms*—Disaggregated RAN, metro-access networks, DU Placement, GNPy, heuristic algorithms, Network Planning**

I. Introduction

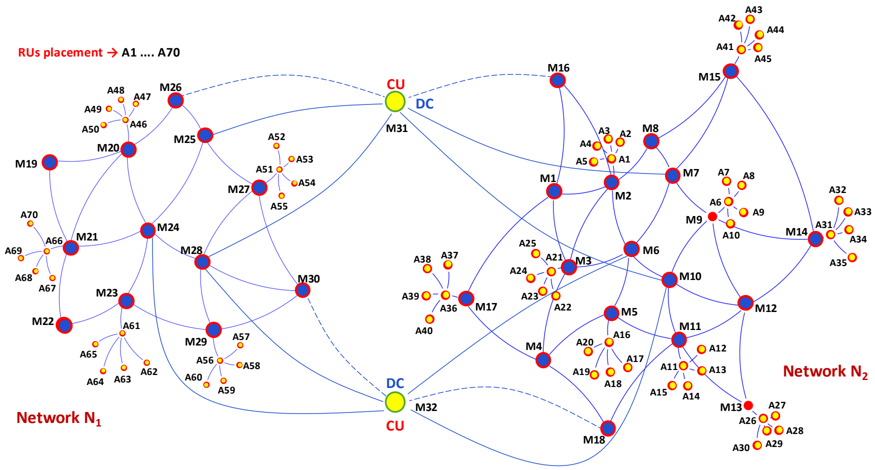
The rapid growth of mobile data traffic, driven by emerging applications such as immersive extended reality, industrial automation, and AI-native services, is placing unprecedented pressure on radio access and transport networks. Fifth-generation (5G) and beyond-5G systems are expected to simultaneously support enhanced mobile broadband (eMBB), ultra-reliable low-latency communications (URLLC), and massive machine-type communications (mMTC), each imposing stringent requirements on latency, capacity, and reliability [1–4]. A key architectural transformation enabling this evolution is the disaggregation of the RAN. This paradigm, driven by virtualization and softwarization, allows network functions to be deployed over general-purpose computing platforms, enabling programmability, scalability, and service-oriented operation [5, 6]. Such disaggregation is at the core of emerging architectures such as cloud-RAN and O-RAN, which aim to promote openness, interoperability, and intelligence in next-generation mobile networks [7, 8]. Moreover, the ability to flexibly distribute functionalities across distributed entities opens the possibility for dynamic network operation and adaptive functional split configurations [9].

However, this architectural flexibility introduces stringent constraints on the transport network. Once RAN functions are distributed, the fronthaul (FH, RU–DU) and midhaul (MH, DU–CU) segments must satisfy strict latency and quality-of-service (QoS) requirements, supported by coherent optical transmission based on DP-16QAM modulation with 64 Gbaud. For example, lower-layer functional splits such as split 7.2 impose tight one-way latency limits on the order of hundreds of microseconds, while also requiring reliable transmission under physical-layer impairments [10, 11]. Furthermore, bandwidth requirements for traditional FH interfaces such as CPRI can become prohibitive, motivating the adoption of functional splits and more efficient transport solutions [10]. These constraints significantly restrict the feasible geographical placement of DU and CU entities and directly impact the design of the transport network.

To address these challenges, converged metro-access optical networks have emerged as a promising platform for supporting disaggregated RAN transport. By leveraging shared optical infrastructure, these networks enable efficient aggregation of access traffic while providing the capacity and reach required for x-haul connectivity [2, 12–14]. At the same time, the feasibility of RU–DU–CU (E2E, end-to-end) connectivity is fundamentally governed by physical-layer impairments as well as strict latency constraints. Therefore, transport feasibility must be evaluated jointly in terms of bit-error-rate (BER) and latency, rather than considering connectivity alone.

In this context, the placement of DU and CU functionalities has been widely studied in the literature. Existing works have investigated dynamic functional split and virtualization as mechanisms to improve flexibility and resource utilization [9], as well as optimization of O-CU and O-DU placement using mathematical programming, heuristics, or machine learning techniques [15, 16]. Other studies have focused on energy-efficient functional split selection and baseband placement under capacity and delay constraints [17, 18]. While these approaches provide valuable insights into architectural flexibility and computational optimization, they typically rely on abstracted transport models and do not explicitly account for detailed physical-layer feasibility in optical networks.

This limitation is particularly critical in converged metro-access environments, where the feasibility of E2E path depends on route-specific optical impairments and latency budgets. Consequently, DU placement cannot be treated as a



Topology Parameters		
Parameter	**N_1**	**N_2**
Total optical nodes	44	63
Metro nodes	12	18
Access nodes	25	45
Avg. node degree	3.42	4.06
Avg. link length [km]	6.9	7.3
Max. span [km]	10	12
Max. hop count (RU–CU)	8	

Fig. 1: Simulated converged metro-access network topology and corresponding parameters.

purely logical or cloud-level optimization problem, but must be addressed jointly with transport-layer feasibility constraints. This observation motivates a physical-layer-aware formulation of the DU placement problem, in which potential locations are evaluated based on their ability to support feasible optical routes under realistic network conditions.

In this work, we investigate DU placements in a converged metro-access optical network under joint BER and latency constraints. We develop a physical-layer-aware evaluation framework that determines the feasibility of E2E paths based on route-level optical performance. The DU placement problem is then formulated as an optimization task aimed at maximizing the number of feasible RU–CU pairs supporting 22.2 Gbps per RU consistent with split 7.2 [19]. To solve this problem, we compare exhaustive search, greedy heuristics (G-Coverage, Degree, Access, and Composite), and genetic algorithm (GA) based approaches, quantifying the trade-offs between optimality and computational complexity. The results provide practical insights into DU deployment strategies for enabling flexible and reliable x-haul transport in disaggregated RAN architectures.

II. Modeling and Simulation Framework

The performance evaluation is carried out on the simulated converged metro-access optical network illustrated in Fig. 1. The topology consists of two interconnected metro-access subnetworks, N_1 and N_2, comprising metro and access segments, reflecting a realistic operator-grade deployment as in [20]. Access nodes are represented by A nodes serving RUs, while metro nodes (M nodes) are modeled as reconfigurable Optical Add-Drop Multiplexers (ROADMs) acting as aggregation points within the optical transport layer. Each access segment consists of four access nodes per cluster that are connected to RUs responsible for E2E traffic. These access nodes are connected to the metro network through circulators, a splitter, and an ROADM, allowing to go seamless from single-fiber BiDi transmission on the access segment to the metro fiber pairs. DU functionalities are assumed to be hosted at selected metro nodes within the network, whereas CU functions are being hosted at centralized data centers (DCs) located at nodes M31 and M32.

The physical-layer performance of each candidate route is evaluated using a Python-based optical network simulation framework build upon GNPy, an open-source QoT estimation tool [21, 22]. This enables realistic estimation of physical-layer impairments such as fiber attenuation, amplified spontaneous emission (ASE) noise, and nonlinear interference (NLI), along multi-span metro-access paths. All fiber spans are modeled as standard single-mode fiber (SSMF) with an attenuation coefficient of 0.25 dB/km, while erbium-doped fiber amplifiers (EDFAs) are assumed with a nominal noise figure of 6 dB [23].

For each route, the effective signal-to-noise ratio (SNR) and the corresponding pre-FEC BER is computed in [20]. Latency is evaluated independently for the FH and MH segments by mapping the physical path length to propagation delay, assuming a propagation delay of approximately $5\,\mu\text{s}\,\text{km}^{-1}$, which represents the dominant latency component in transparent optical transmission. An evaluated path is considered feasible only if it simultaneously satisfies BER and latency constraints, namely BER $\leq 10^{-2}$, FH latency $\leq 250\,\mu\text{s}$, and MH latency $\leq 1\,\text{ms}$, consistent with 3GPP split 7.2 requirements [10, 20, 23, 24].

Based on this physical-layer evaluation, a feasibility matrix is precomputed for all path combinations. This matrix encodes whether a given RU can be served through a specific DU under the imposed constraints and forms the basis for the DU placement optimization. The placement problem is formulated as selecting a subset of K metro nodes that maximizes the number of feasible RU–CU pairs, while also considering the number of feasible route instances as a secondary metric to capture route diversity. We study three deployment strategies: exhaustive search, greedy selection, and a GA. Exhaustive search is used as an optimal benchmark. The greedy approach incrementally selects DUs that provide the maximum coverage gain, while the GA enables scalable near-optimal solutions for larger search spaces. The overall workflow is summarized in Algorithm 1, which consists of two phases. In Phase 1, feasibility is evaluated by constructing E2E paths under hop

Algorithm 1 Physical-layer-aware DU placement framework

Input: Topology G, RU set $\mathcal{R}$, DU candidates $\mathcal{D}$, CU set $\mathcal{C}$, budget K, hop limit H, mode
Output: Selected DU set $\mathcal{D}_K^\star$
Phase 1: Feasibility Evaluation
for all $(r, d, c) \in \mathcal{R} \times \mathcal{D} \times \mathcal{C}$ **do**
  Construct RU–DU–CU path with hop constraint H
  if path exists **then**
    Compute GSNR, SNR, BER, and FH/MH latency
    Set $\mathbf{F}(r, d, c) \leftarrow 1$ if constraints are satisfied, else 0
  end if
end for
Phase 2: DU Placement Optimization
if mode = Exhaustive **then**
  for all $\mathcal{S} \subseteq \mathcal{D}$ such that $|\mathcal{S}| = K$ **do**
    Evaluate coverage using $\mathbf{F}$ (tie-break: routes)
  end for
  $\mathcal{D}_K^\star \leftarrow$ best $\mathcal{S}$
else if mode = Greedy-Coverage **then**
  Select DU maximizing marginal RU–CU coverage
else if mode = Greedy-Degree **then**
  Select DU with highest node degree
else if mode = Greedy-Access **then**
  Select DU closest to maximum number of access nodes
else if mode = Greedy-Composite **then**
  Select DU maximizing weighted score (degree, access proximity, distance)
else if mode = Genetic **then**
  Initialize population with K-active DU chromosomes
  while termination criterion not met **do**
    Evaluate fitness (coverage with route tie-break)
    Apply selection, crossover, mutation, and repair
  end while
  $\mathcal{D}_K^\star \leftarrow$ best chromosome
end if
return $\mathcal{D}_K^\star$

constraints and verifying physical-layer and latency requirements. In Phase 2, DU placement is optimized according to the selected mode, which determines the approach (exhaustive, greedy, or GA). The framework ultimately outputs the set of selected DUs that maximizes E2E coverage while satisfying all feasibility constraints.

III. Results and Discussions

Fig. 2 shows the standalone coverage of individual DU locations, i.e., the number of RU–CU pairs served when a single DU is deployed on a node. A significant variability is observed across the network: centrally located nodes (e.g., M6, M10, M11, ...) serve up to a maximum of 72 pairs, whereas other nodes exhibit substantially lower coverage. This behavior reflects the joint impact of node connectivity and physical-layer constraints. Nodes with high coverage benefit from favorable connectivity to both access clusters and CU locations, as well as shorter propagation distances. In contrast, other nodes are limited by increased latency and degraded signal quality along longer routes. The maximum standalone coverage ranges from ~21% to ~64% of all RU–CU pairs, indicating that single-DU deployment is insufficient and motivating multi-DU placement optimization. This also explains why high-performing nodes (e.g., M6) are consistently selected in the initial stages of the algorithms. Overall, the results highlight

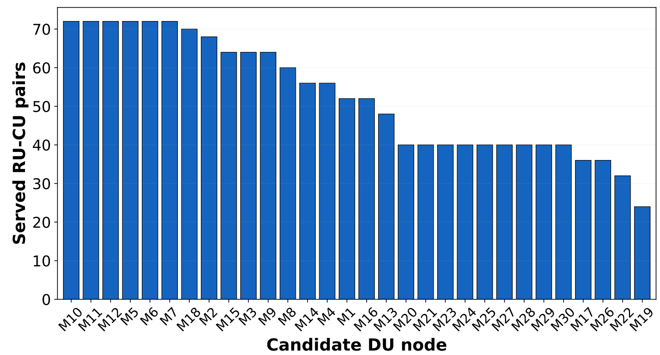


Fig. 2: Standalone DU coverage across metro nodes under feasibility constraints.

that DU effectiveness is governed by physical-layer feasibility rather than purely topological properties.

Fig. 3 presents the number of served RU–CU pairs as a function of the number of deployed DUs (K). All approaches reach optimal coverage (112 pairs) at a maximum of $K = 4$. Greedy-Coverage attains the same optimal solution as exhaustive search, selecting nodes M6 and M24, demonstrating that a simple heuristic can identify the global optimum. Greedy-Access and Greedy-Composite show similar behavior, while Greedy-Degree requires $K = 3$, indicating that degree-based metrics alone are insufficient to capture physical-layer feasibility. The GA also converges to the optimal solution, albeit with higher complexity. Beyond $K = 2$, no further increase in coverage is observed, implying that additional DUs contribute only to route diversity rather than connectivity.

Fig. 4 shows the total number of E2E feasible routes versus K. The number of feasible routes increases steadily with K, from $\sim 3.5 \times 10^3$ at $K = 2$ to nearly 1.4×10^4 at $K = 10$ under optimal placements. This indicates that, while a small number of DUs ensure full connectivity, additional DUs significantly enhance route diversity, enabling improved resilience, load balancing, and service flexibility. Greedy-Coverage closely matches exhaustive and GA performance across all K, whereas Greedy-Degree consistently underperforms. Greedy-Access and Greedy-Composite achieve full coverage but yield fewer routes, reflecting locally optimal but globally suboptimal placements. These results demonstrate that, once coverage is satisfied, route diversity becomes the

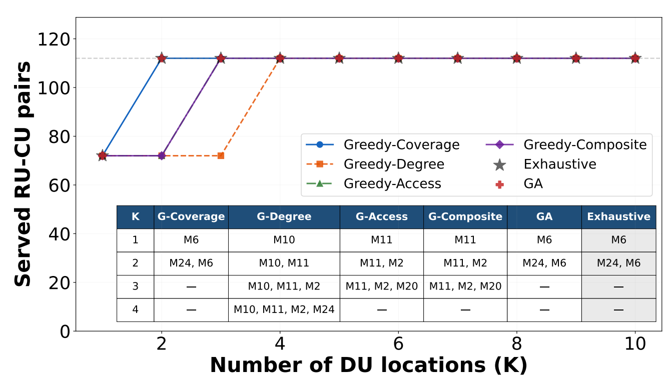


Fig. 3: RU–CU coverage vs. number of DU locations (K).

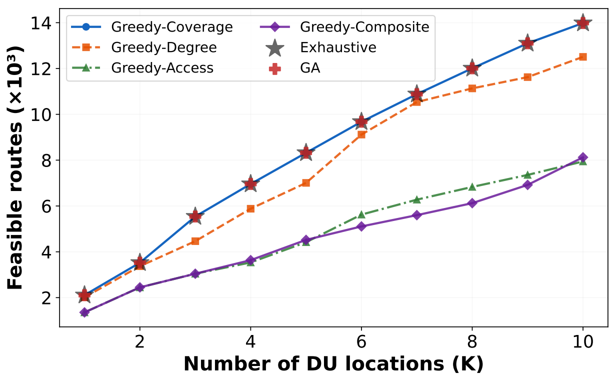


Fig. 4: Feasible RU–DU–CU routes vs. K, showing increasing route diversity.

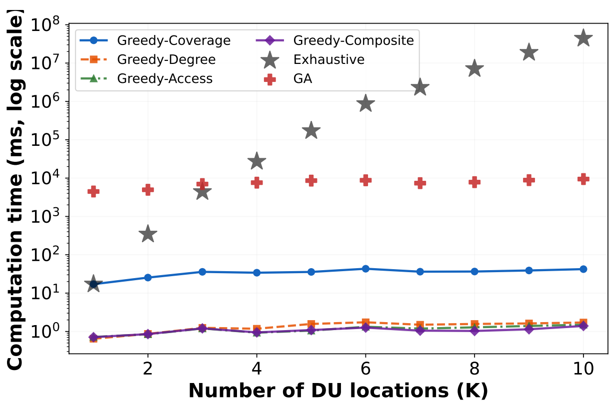


Fig. 5: Computation time vs. K (log scale) for different placement strategies.

dominant optimization objective.

Fig. 5 reports the computational complexity of the evaluated strategies. Exhaustive search exhibits exponential growth, increasing from milliseconds at smaller K to $\sim 10^8$ ms at $K = 10$, making it computationally expensive for larger instances due to the combinatorial search space. In contrast, all greedy variants maintain near-constant execution time in the millisecond range, demonstrating excellent scalability. GA provides a compromise, achieving optimal solutions with execution times in the order of seconds. Overall, Greedy-Coverage achieves optimal performance in both coverage and route diversity while maintaining several orders of magnitude lower complexity compared to exhaustive search, making it the most practical solution for large-scale deployments.

IV. Concluding Remarks

This paper presented a physical-layer-aware framework for the number of DU required and their optimal placement in converged metro-access optical networks under joint latency and QoT constraints. By leveraging a GNPy-based QoT-aware simulation framework, the feasibility of E2E routes was evaluated considering BER and strict FH/MH latency requirements. The results demonstrate that full RU–CU coverage can be achieved with a minimal number of DU locations (e.g., $K = 2$) in our scenario, highlighting the strong influence of network topology (nodes connectivity) and physical-layer impairments on feasible connectivity. However, beyond this point, the optimization objective shifts from coverage to route diversity, which becomes critical for improving network resilience, flexibility, and load balancing capabilities.

A comparative analysis of exhaustive, greedy, and genetic algorithms shows that simple greedy strategies, particularly the coverage-based heuristic, are able to achieve optimal solutions with negligible computational complexity. These findings indicate that physically-aware G-Coverage and GA approaches provide an effective and scalable solution for DU placement in practical deployments. The proposed framework establishes a foundation for future extensions toward traffic-aware and capacity-driven optimization, enabling dynamic and service-oriented resource provisioning in next-generation disaggregated RAN architectures.

Acknowledgments: Ahtisham Ali's work is fully funded by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie project EWOC under GA No. 101073265. This work has also received support from the MSCA project NESTOR under GA No. 101119983, the EU Horizon Project ALLEGRO GA No. 101092766, the PNRR-NGEU (MUR–DM117/2023), and the EU NGEU partnership on 'Telecommun. of the Future' (PE00000001 – program 'RESTART').

References

[1] Z. Vujicic *et al.*, "Towards virtualized optical-wireless heterogeneous networks," *IEEE access*, 2024.
[2] F. B. Saghezchi *et al.*, "Drive towards 6G," in *Enabling 6G Mobile Networks*. Springer, 2021, pp. 3–35.
[3] A. Ghosh *et al.*, "5G evolution: A view on 5G cellular technology beyond 3GPP release 15," *IEEE Access*, vol. 7, pp. 127 639–127 651, 2019.
[4] ITU, "IMT Vision - Framework and overall objectives of the future development of IMT for 2020 and beyond," ITU-R, Tech. Rep. M.2083-0, 2015.
[5] L. M. P. Larsen *et al.*, "A Survey of the Functional Splits Proposed for 5G Mobile Crosshaul Networks," *IEEE Commun. Surveys & Tutorials*, vol. 21, no. 1, pp. 146–172, 2019.
[6] L. Bonati *et al.*, "Intelligence and learning in O-RAN for data-driven NextG cellular networks," *IEEE Commun. Magazine*, vol. 59, no. 10, pp. 21–27, 2021.
[7] A. Garcia-Saavedra and X. Costa-Perez, "O-RAN: Disrupting the virtualized RAN ecosystem," *IEEE Commun. Standards Magazine*, vol. 5, no. 4, pp. 96–103, 2021.
[8] G. O-RAN Alliance, Alfter, "O-RAN WhitePaper—Building the Next Generation RAN," Tech. Rep., October 2018.
[9] R. J. B. Pousa *et al.*, "Road to Dynamic Functional Split in Radio Access Networks," in *Proc. IEEE MELECON*, 2024, pp. 844–849.
[10] 3GPP, "Study on CU-DU lower layer split for NR; (Release 15)," 3GPP, Tech. Rep. TR 38.816 V15.0.0 (2017-12), 2017.
[11] A. Tzanakaki *et al.*, "Optical transport networks converging edge compute and central cloud: an enabler for 6G services," in *Proc. OFC*. IEEE, 2024, pp. 1–3.
[12] M. Klinkowski, "Latency-aware DU/CU placement in convergent packet-based 5G fronthaul transport networks," *Applied Sciences*, vol. 10, no. 21, p. 7429, 2020.
[13] G. Rizzelli *et al.*, "Experimental demonstration of in-field 400G coherent metro-access convergence," in *Proc. OFC*. Optica Publishing Group, 2024, pp. W1J–1.
[14] G. Rizzelli *et al.*, "Analysis and experimental demonstration of possible architectures for future coherent metro+ PON converged networks," *JOCN*, vol. 17, no. 2, pp. A142–A154, 2025.
[15] R. Joda *et al.*, "Deep reinforcement learning-based joint user association and CU–DU placement in O-RAN," *IEEE Trans. on Network and Service Management*, vol. 19, no. 4, pp. 4097–4110, 2022.
[16] H. Hojeij *et al.*, "On Flexible Placement of O-CU and O-DU Functionalities in Open-RAN Architecture," *IEEE Trans. on Network and Service Management*, vol. 22, no. 1, pp. 660–674, 2025.
[17] N. Sen and A. F. A, "Towards Energy Efficient Functional Split and Baseband Function Placement for 5G RAN," in *Proc. IEEE NetSoft*, 2023, pp. 237–241.
[18] S. Zeb *et al.*, "Power aware data center placement in WDM optical networks," in *Proc. IEEE CCNC*, 2021, pp. 1–6.
[19] A. Ali *et al.*, "Optimizing RAN X-Haul Performance Through Targeted Hollow-Core Fiber Deployment in Converged Metro-Access Networks," in *Proc. SoftCOM*, 2025, pp. 1–6.
[20] A. Ali *et al.*, "Evaluating Metro-Access Optical Networks as-a-Service for Flexible RAN X-haul," in *Proc. IEEE CSCN*, 2025, pp. 1–6.
[21] Telecom Infra Project, "GNPy: Open Optical Performance Modeling Tool," 2019.
[22] V. Curri, "GNPy model of the physical layer for open and disaggregated optical networking," *JOCN*, vol. 14, no. 6, pp. C92–C104, 2022.
[23] A. Ali *et al.*, "Statistical Analysis of end-to-end Route Feasibility in Converged Metro–Access Optical Networks," *IEEE PTL*, pp. 1–1, 2026.
[24] PROTEUS-6G Consortium, "D2.1: 6G x-haul network requirements," PROTEUS-6G Project, EU Horizon Europe, Tech. Rep., 2024.