



UNIVERSITÀ DEGLI STUDI DI NAPOLI
FEDERICO II

itee^{PhD}
information technology
electrical engineering



DIE
TI

UNI
NA

Alessandro Mauro

End-to-End service optimization in programmable networks

Tutor: Antonia Maria Tulino
co-Tutor: Prof. Jaime Llorca

Cycle: XXXIX

Year: 2024/2025

My background

- M.Sc. Degree in Computer Science
 - Thesis: «*Tecniche di Machine Learning per Lap Time Prediction di veicoli sportivi*», supervised by Prof. Sergio Di Martino
- Ph.D. fellowship founded by PNRR Partenariato Esteso PE14 "RESTART - RESearch and innovation on future Telecommunications systems and networks, to make Italy more smART", Spoke 4

Research Field of Interest

- Main field of interest is **Cloud-Edge Continuum**:
 - End-to-End Service Orchestration for Cloud-Edge Continuum
 - Modeling, analysis, evaluation, and control of next-generation computing networks (e.g., distributed cloud, mobile edge, and fog computing) and services
 - AI for Network
 - Network for AI

Summary of study activities

	Courses	Seminars	Research	Tutorship	Total
Total	22	5.7	43	0.4	71.1
Expected (2nd year)	10- 20	5 - 10	30 - 45	0 – 1.6	

Conferences and seminars attended

- *RESTART Plenary Dissemination Workshop, Poster presentation, Napoli 30 June 2025.*
- *RESTART SUPER and NETWIN Workshop, Poster presentation, Rome 24 September 2025.*
- *RESTART SUPER and NETWIN Workshop, Dissemination Talk, Rome 24 September 2025.*
- *RESTART Spoke 4 Workshop, Dissemination Talk, Torino 30 October 2025.*

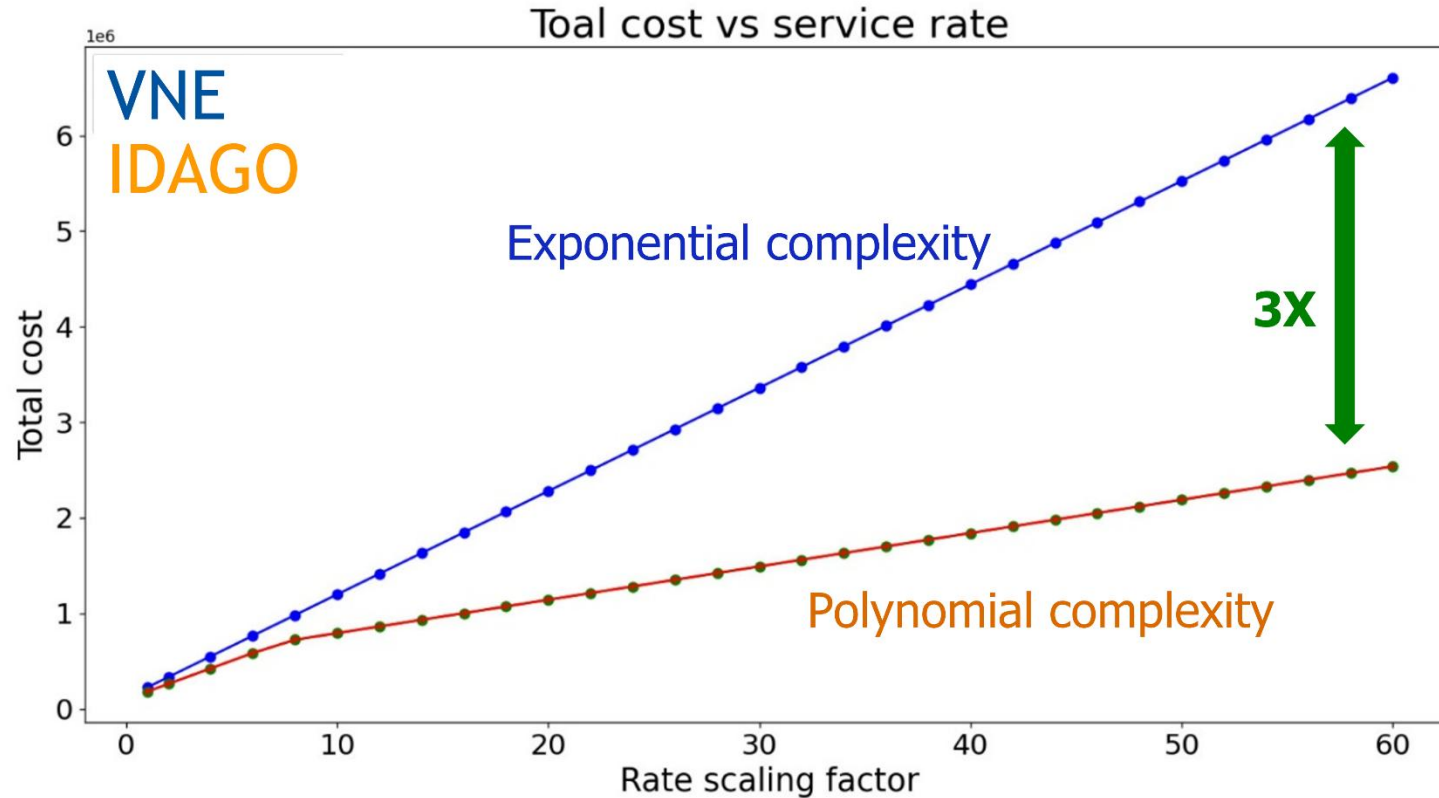
Research area

- **Objective:** optimize the orchestration of application over distributed edge-cloud networks
 - Function placement
 - Flow routing
 - Comm./comp./storage resource allocation

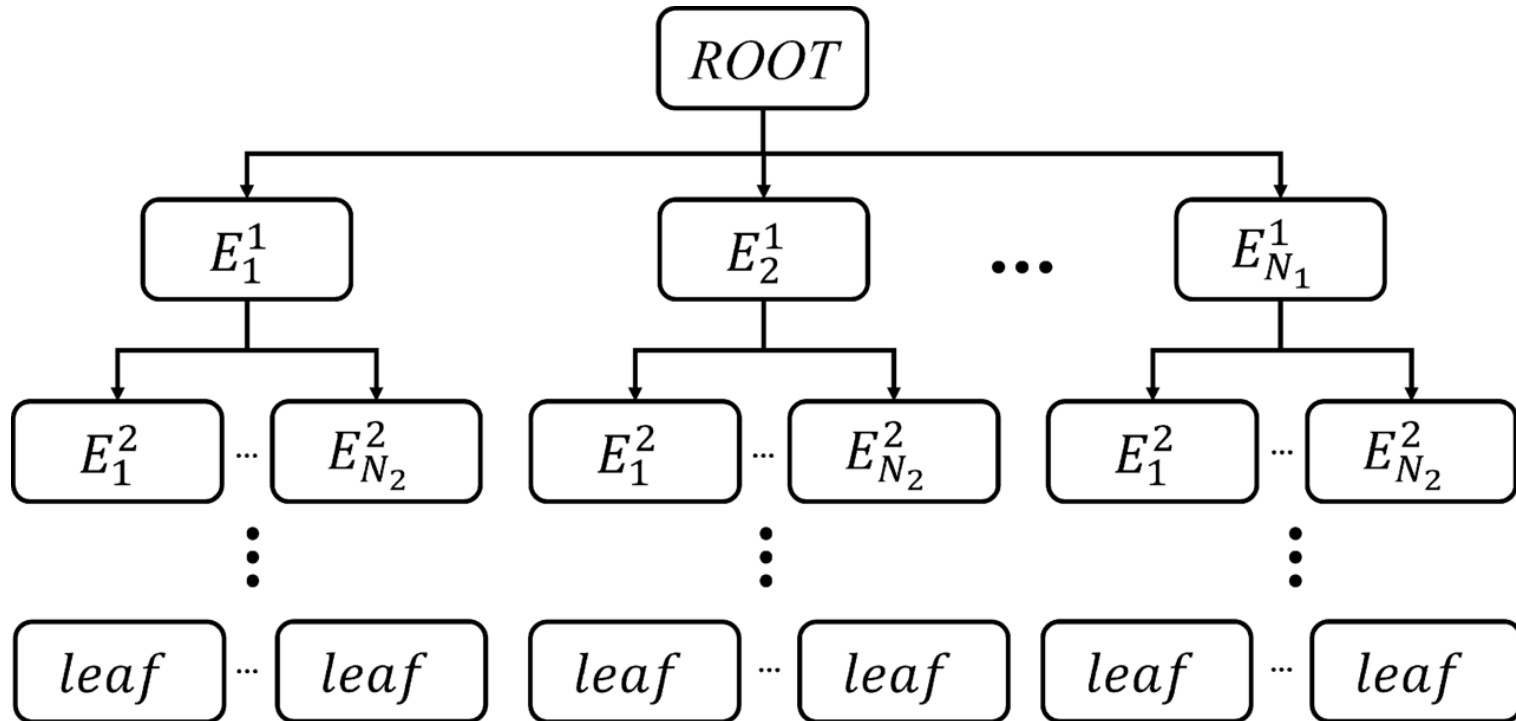
Research results - IDAGO

- Polynomial-time approximation algorithm
 - (multiple) embeddings for the optimal distribution (function/data placement, flow routing, resource allocation) of NextG information-aware DAG services over NextG cloud-integrated networks
 - An embedding is composed of a function mapping, commodity mapping, and an associated probability
- Information-aware DAG-to-Forest service graph transformation procedure
 - Maximizing replication opportunities
 - Adapting LP relaxation, decomposition, and rounding techniques

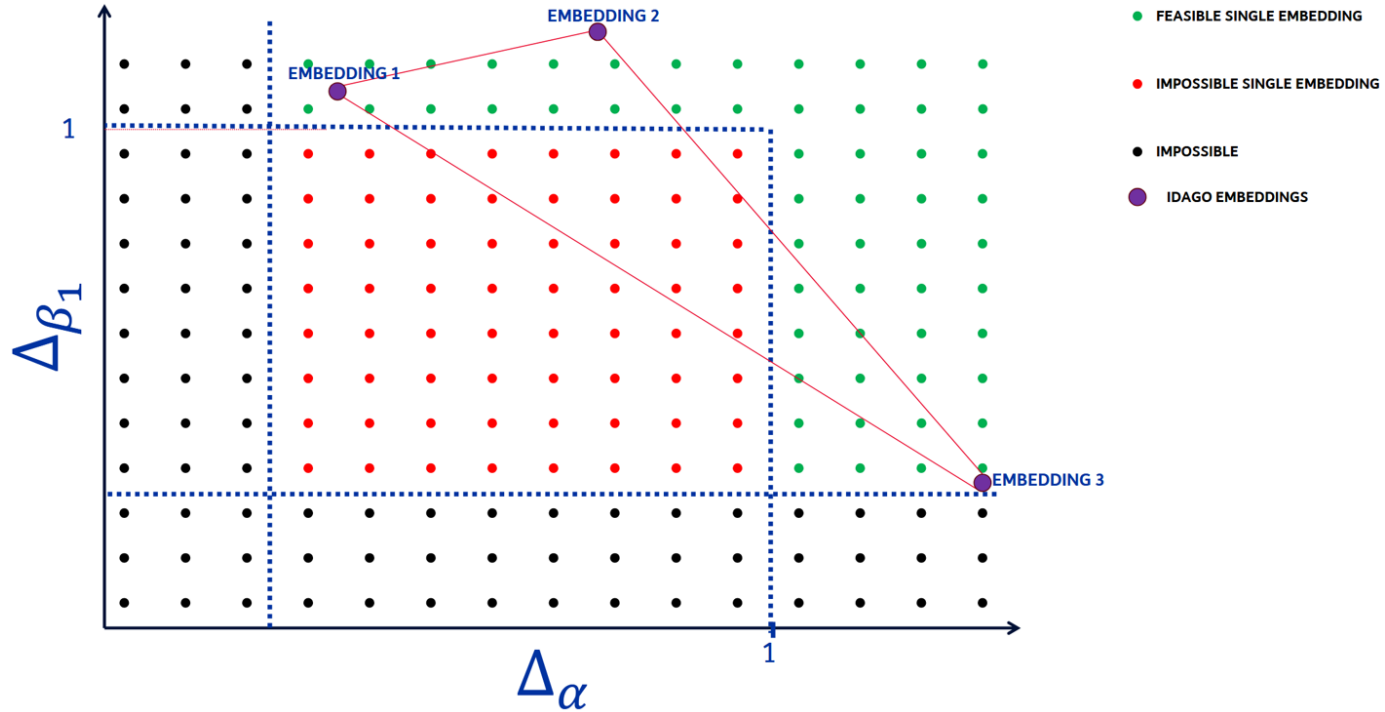
Research results - IDAGO



Research results – IDAGO improvements



Research results – IDAGO improvements



Research results – IDAGO

improvements

- Computing the solution in one shot, leveraging LP solution decomposition but avoiding randomized rounding.
- Constructs the solution in one shot, leveraging new analysis for the probability of violations of constraints, avoiding agnostic LP decomposition.
- Dynamically alternating between multiple approximate solutions over time, leverages spare capacity opportunistically to mitigate the risk of capacity violations.

Research activities

- Development of End-to-End (or long-term) Orchestration algorithms.
- Such algorithm are used in two PoC of RESTART:
 - **SUPER PoC/Demo #6** “Energy-Efficient, End-to-End Orchestration of Next-Generation, AI-Based Video Streaming Applications” [Integration with the SUPER Orchestrator]
 - **Netwin PoC/Demo #3** “Interactive Virtual Assistant with Conversational AI Capabilities”
- Development of dual-timescale optimization solutions (integration of long-term and short-term algorithms):
 - Integrate long-term algorithm with a Multi-Armed Bandit (MAB) algorithm for dynamic resource autoscaling
 - Integrate long-term algorithm with reinforcement learning (RL)-based algorithms for adaptive routing decisions

Plan for year three

- Expand topics I have start exploring:
 - Use of Behavior Trees to capture applications' runtime control logic to characterize the functional relationships in the service graphs used as input for orchestration algorithms
 - Finalize the development of dual-time orchestration solutions
 - Finalize PoC activities in preparation for presentation at the final RESTART Plenary
- Period abroad
 - Centre Tecnologic de Telecomunicacions de Catalunya (CTTC/CERCA), Castelldefels, Spain, supervised by Prof. Jaime Llorca
- PhD Thesis
 - AI for networks: AI techniques to optimize service orchestration
 - Networks for AI: designing and testing orchestration solutions tailored for AI-based applications.

Research products

[P1]	Mauro, A., Tulino, A. M., & Llorca, J. (2025). End-to-End Orchestration of NextG Media Services over the Distributed Compute Continuum. IEEE Transactions on Mobile Computing. Status: Early Access.
[P2]	Mauro, A., Origlia, A., Tulino, A. M., & Llorca, J. (2025). DORIAN: Deterministic Orchestration of AI-Centric Applications over Distributed Edge-Cloud Networks. Computer Networks. Status: Under Review.
[P3]	Mauro, A., Tulino, A. M., & Llorca, J. (2025) Robust and Predictable Orchestration of Distributed Multiuser AI-Powered Applications. ICC 2026 - IEEE International Conference on Communications: Next-Generation Networking & Internet. Status: Submitted.
[P4]	Di Bratto, M., Origlia, A., Llorca, J., Detti, A., Mauro, A. , Grazioso, M., Vitale, V. N., Russo, V., Mancini, A., Perrino, A., Napolitano, N., Della Corte, G., Tulino, A. M., Piane, S., & Tennirelli, M. (2025). Automatic positioning of AI microservices on NextG networks to support interactive holograms. Ital-IA 2025. Status: Accepted.
[P5]	Poster presentation, RESTART Plenary Dissemination Workshop, Napoli 30 June 2025.
[P6]	Poster presentation, RESTART SUPER and NETWIN Workshop, Rome 24 September 2025.