

# Integrating Artificial Intelligence Techniques for the Management of Softwarized Networks

Ph.D. Thesis Summary

**Candidate:** Antonino Angi

**Advisor:** Guido Marchetto

Nowadays, the complexity of networks is continuously increasing due to the growing number of connected devices and a greater demand for data-intensive applications that often require low latency and high throughput. These demands pose critical challenges for complex and highly connected topologies, such as data centers, where big volumes of traffic data need to be processed and transferred at fast rates. Alongside this, traditional static network architectures often lack the flexibility to adapt to dynamic traffic patterns or support real-time management, resulting in degraded performance, increased latency, and potential bottlenecks.

An approach to support these challenges is provided by the concept of Software-Defined Networking (SDN), which simplifies network deployment and enables real-time adjustments, ensuring automated adaptability to dynamic network conditions. In this context, recent management approaches for softwarized networks leverage Artificial Intelligence (AI) and Machine Learning (ML) techniques to analyze traffic conditions, aiming to further reduce network reaction time and improve the accuracy of automated decisions. To intersect these advantages, data-plane programming has been introduced to enable fine-grained control over packet processing, allowing more efficient packet customization at the hardware level. However, while these techniques have shown great capabilities in improving network performance, they still lack structured mechanisms for efficiently distributing traffic loads in critical scenarios (*e.g.*, network congestion) or for appropriately integrating AI/ML methods that, due to their resource-intensive nature, can negatively impact the network's forwarding efficiency.

In this thesis, we present SDN solutions that leverage AI/ML methods and data-plane programmability to enable intelligent network routing for dynamic traffic demands on complex topologies, such as data centers. We begin by introducing Howdah, a profiling mechanism that, by performing traffic classification, offers adaptation to different traffic classes under different network congestion scenarios. Building on this, we propose two advanced approaches for more autonomous traffic management: ROAR and ART. The former employs a Multi-Agent Reinforcement Learning (MARL) mechanism to implement adaptive routing policies based on real-time network conditions. By embracing the In-Network ML paradigm, the latter translates Deep Reinforcement Learning

(DRL) decisions into lighter Decision Trees (DT), offering a more efficient solution that can be implemented directly inside switches without requiring external units.

Nevertheless, while network programming provides this high degree of customization, it is also known to be challenging, even for experts, requiring trial-and-error approaches with iterative debugging and log file analysis. This process is not only time-consuming but also prone to errors, slowing down the entire network customization process. To address this challenge, in this thesis, we also explore the Intent-based Networking (IBN) paradigm, which simplifies network programming by allowing programmers to specify high-level objectives (*i.e.*, intents) and have them translated into low-level network configurations. Thus, we propose two separate solutions: NLP4 and NAIL. The former combines Natural Language Processing (NLP) techniques with a MultiLayer Perceptron (MLP) model to translate high-level intents into mid-level policies, which are then integrated into network configurations through an API. The latter is a network transpiler architecture that adopts NLP and a specialized API to insert, monitor, and remove intents from a network without restarting the system.

In conclusion, we believe the proposed solutions contribute to enhancing programmable networks, addressing the growing complexity that modern networks face by providing greater flexibility and adaptation to dynamically changing conditions.