Orchestrating Edge Computing Services with Efficient Data Planes

Doctoral Program in Computer and Control Engineering (36thcycle)

Candidate: **Federico Parola**

Supervisor: Prof. Fulvio Risso

In the rapidly evolving landscape of telecommunications, the advent of the 5G technology has emerged as a revolutionary force, promising to redefine the way we connect, communicate, and experience the digital world. One of the architectural cornerstones enabling 5G innovation is represented by Multi-access Edge Computing (MEC), which involves deploying computing resources at the edge of the network, closer to the end-users and devices, guaranteeing lower latencies and reducing the amount of data that needs to traverse the network backbone. This paradigm shift is coupled with the almost complete migration to Network Function Virtualization (NFV) for networking tasks which, by moving traditional networking functions from dedicated hardware appliances to software-based solutions, provides greater flexibility, scalability, and efficiency in the deployment and operation of network services.

The transfer of an increasing share of computing workloads to small edge deployments, coupled with the execution of network functions on the same general-purpose servers, pose non-trivial challenges in how edge data centers are managed. This makes resource consumption, an already strategic topic in the cloud context, an even more crucial requisite to make edge computing an effective solution. Hence, in this dissertation, we analyze the problem of resource optimization in edge data centers, with the goal of maximizing the number of applications and network functions that can be deployed in the constrained network edge and benefit from its improved performance. We do so by focusing on three main optimization areas. We start by studying how to enable the **consolidation of cloud and network workloads on the same physical server** as the first building block for a more efficient and flexible edge. Since kernel-bypass frameworks widely used for packet processing (e.g., DPDK) pose non-negligible drawbacks in terms of integration and resource-sharing with traditional kernel-based cloud applications, we explore the capabilities recently introduced by the eBPF and XDP technologies to perform fast packet processing in the kernel, as well as to flexibly redirect packets to user space without completely bypassing the former. We show how, thanks to these technologies, a combination of in-kernel and user space packet processing can be leveraged to maximize the performance of both cloud native

and network workloads. Then, we move to **enabling secure and efficient multi-tenancy** on edge servers, aiming to facilitate safe access to this precious resource. To this end, we design a virtualization environment based on unikernels, to combine the leanness of containers with the strong isolation of virtual machines, and integrate it with a high-performance zero-copy data plane to handle inter-service communication. Finally, we focus on the increasing demands that interconnecting application modules poses on data center networking, both in terms of performance and flexibility. In this respect, we design ways of **improving the management of services through eBPF**, for both traditional cloud workloads and chains of network functions. Focusing on Kubernetes, the most widespread cloud orchestrator, we design an alternative to classic, monolithic eBPF-based networking providers, proposing an approach based on modular, reusable building blocks coming from the traditional networking world, such as routers, bridges, and load balancers, which despite its higher level or abstraction does not compromise performance. We also enable applying the automatic scaling capabilities of Kubernetes to network function chains, leveraging a mechanism of flexible cross-connections to efficiently interconnect replicas of network functions.

Overall, the work presented in this dissertation contributes to enhancing the capabilities of edge data centers to support the ever-increasing load they have to handle, with a high degree of efficiency, making the promises of 5G closer to becoming reality.