Toward a Cloud Continuum for Efficient 6G Infrastructures

Jacopo Marino

Edge computing, 5G/6G networks, and containerization have created the conditions for a new paradigm: the edge-to-cloud continuum, in which heterogeneous infrastructures spanning far-edge devices, telco facilities, and cloud data centers are treated as a unified pool of resources. Despite this vision, deployments remain fragmented. Within each fragment, resources are managed through container orchestrators, with Kubernetes as the de facto standard; yet each cluster has an independent control plane, exposing only a partial view of available resources, and applications must be assigned to specific infrastructures at deployment, preventing transparent migration, dynamic load redistribution, and cross-domain sharing.

This thesis lays the foundations of a computing continuum across distributed cloud infrastructures. Part I builds the architectural foundations, while Part II validate them through distributed robotics. The parts share a goal but address distinct questions, and can largely be read independently.

On the infrastructure side, the thesis first examines the physical substrate available to telco operators. Drawing on publicly available telco data, it shows that operators already possess tens of thousands of central offices, telephone exchanges, and 5G cell sites suitable for hosting edge computing resources. A survey of existing multi-cluster technologies identifies Liqo as the most comprehensive production-ready foundation, while highlighting the open challenges that current solutions fail to address.

Building on this analysis, the thesis presents the Flexible, scaLable, secUre, and decentralIseD Operating System (FLUIDOS), which realizes the continuum through three transparency properties: deployment, communication, and resource availability. Central to FLUIDOS is the REsource Advertisement and Reservation (REAR) protocol, enabling standardized, efficient, and secure resource discovery and allocation across heterogeneous domains. REAR supports a rich data model covering heterogeneous resource types, with authentication grounded in decentralized identifiers and verifiable credentials. Complementing these foundations, the thesis proposes infrastructure-level data spaces that

leverage multi-cluster capabilities to enable controlled cross-domain data sharing while preserving data sovereignty and regulatory compliance, with negligible overhead.

The continuum principles are then applied to distributed robotic systems, where latency, reliability, and energy efficiency constraints provide a demanding validation context. Three contributions are presented. The Scalable Middleware 2 (SM2) decouples local inter-process communication from network transport, preventing the performance degradation observed in DDS-based ROS2 configurations under adverse conditions, and reduces CPU usage by 3× and memory consumption by 4× compared to state-of-the-art alternatives. Four switching architectures for migrating stateless ROS2 computation between Kubernetes clusters are evaluated, showing that a Lifecycle Node-based approach achieves switching times at least 85% faster than redeployment-based alternatives. Finally, a Multiplexer component is proposed to switch between local and remote navigation streams based on message freshness and reliability, ensuring continuous and safe robot operation even under severe packet loss: at 55% packet loss, the approach achieves a 31% improvement in navigation performance over remote-only control, while a complementary Lifecycle Orchestrator reduces onboard CPU usage by up to 99% during periods of stable remote connectivity.

Taken together, these contributions establish a coherent architectural and experimental foundation for the computing continuum, spanning its physical motivations, enabling protocols, data governance mechanisms, and application to real-world robotic scenarios, demonstrating how cloud continuum architectures can serve as a concrete enabler for the efficient deployment of 6G network services.