International Collaborative Experiment for RL-based Multi-Resource Management on COSMIC Testbed

Arslan Qadeer[1], Myung J. Lee[1] Daiki Nobayashi[2], and Kazuya Tsukamoto[2]
[1]Department of Electrical Engineering, The City College of New York, City University of New York, USA
[2]Graduate School of Engineering/Computer Science and Systems Engineering, Kyushu Institute of Technology, JPN

Abstract

To brace next-generation latency and capacity sensitive and diverse mobile applications for 6G and IoT across network domains, there is a need to co-develop high-speed communications and efficient computing infrastructure. Edge Cloud (EC) is poised to address these constraints by providing compute resources at the edge of the network. Yet, the EC being regionally domestic with a smaller scale, faces the challenges radio access bandwidth and computational throughput, calling for a new SDN EC system architecture integrated with the core cloud for resource allocation and task scheduling. Such architecture will promote resource sharing among multiple and heterogeneous network domains. The NSF IRNC supported COSMIC (COSMOS Interconnecting Continents) testbed proposes to build a fully programmable network (from the optical and radio physical layers and above) and computational infrastructure (edge clouds). In the proposed COSMIC architecture, data center near to the mobile users can serve as an edge cloud and be used to offload service requests of real-time and latency-sensitive applications. The data center in a remote LAB or other continents will be emulated as a core cloud, and the delay-tolerant services can be offloaded to the data center to achieve efficient resource sharing in a multi-domain network environment. The testbed is well suited to such category of experiments which involve emulated mobility between multiple edge network domains, potentially in different continents. The COSMIC's global sites will ensure a worst-case evaluation environment for stress-testing of real-time applications. Diverse quality of service requirements germane to IoT traffics will be managed by the proposed SDN architecture. To cope with real-time constraint and scalability issues, Reinforcement Learning based approach is proposed to optimally manage wireless access bandwidth, flow route resource, and VM resources of edge and core clouds. Moreover, Collaborative Learning based approach (such as Federated Learning) is also proposed to flexibly address critical issues such as data privacy, data access to heterogeneous data among multiple edge networks with diverse characteristics. A use-case scenario like edge cloud client environment for business travelers across multiple continents will be developed over the globally connected COSMIC testbed.

While COSMIC testbed development progressing, the RL-based multi-resource management model for applications with bounded latency will first be developed and experimented over NICT StarBED testbed (a large scale Edge Clouds emulation testbed built on SDN testbed RISE by NICT) via CCNY-Kyutech testbed connection supported by NSF/NICT JUNO2. CCNY-Kyutech testbed will also be connected to COSMOS testbed. The model will also be developed and tested over NICT Smithsonian testbed (a large scale Simulation and Emulation Federation Technology on StarBED) supported by NICT-Kyutech Matching Project.