BeGREEN Intelligent Plane for AI-driven Energy Efficient O-RAN management

Miguel Catalan-Cid, Jorge Pueyo
i2CAT Foundation, Spain
{miguel.catalan, jorge.pueyo}@i2cat.net

Juan Sanchez-Gonzalez
Universitat Politècnica de Catalunya, Spain
juansanchez@tsc.upc.edu

Jesus Gutierrez
IHP - Leibniz-Institut für innovative Mikroelektronik, Germany
teran@ihp-microelectronics.com

Mir Ghoraishi
Gigasys Solutions Ltd, UK
mir@gigasys.co.uk

***Abstract*—Cellular networks are undergoing a revolutionary transform with the advent of O-RAN architectures and AI/ML solutions. O-RAN's Non-Real-Time and Near-Real Time RAN Intelligent Controllers open the door to the implementation of automated control-loops that can provide RAN optimisations in numerous scenarios and use cases, and which can be further empowered by AI-driven approaches. Energetic sustainability has raised as one of the main optimisations targets due to the impact of mobile networks on global energy consumption. To this end, the BeGREEN project aims at enhancing the energy efficiency of beyond 5G networks by defining novel AI/ML-based methods at RAN and edge infrastructure. This paper presents BeGREEN Intelligent Plane, a novel framework which implements and exposes AI/ML workflows to O-RAN-based optimisations targeting energy efficiency. We also describe an exemplary application of the Intelligent Plane and its AI Engine, which aims at providing AI-driven cell on/off control.**

I. Introduction

The transition from 5G to beyond 5G (B5G) and 6G mobile communication networks brings a paradigm shift not only in terms of enhanced performance and increased connectivity, but also in addressing critical issues related to the environmental implications associated to a higher energy consumption. Improving the planning, deployment, and management of B5G and 6G networks is imperative to counteract the rising energy consumption trend. Overcoming these challenges require innovative architectural revisions and novel algorithmic solutions to promote sustainability and mitigate the environmental impact of cellular networks [1]. Notably, the Radio Access Network (RAN) consumes more than 70% of the total energy of a 5G system, making its optimisation a top priority.

The consolidation of the O-RAN architecture, which advocates for disaggregated, virtualized and software-based components, connected through open and standardised interfaces, entails a significant opportunity to intelligently manage the RAN with the aim of improving the network performance, and reduce energy consumption [2]. Particularly, the Non-Real-Time RAN Intelligent Controller (non-RT RIC), and Near-Real Time RAN Intelligent Controller (Near-RT RIC) provide the required functionalities to develop and host the so-called rApps and xApps implementing, respectively, long-term and almost real-time optimisations through automated control-loops.

Furthermore, the integration of Artificial Intelligence and Machine Learning (AI/ML) introduces a cognitive layer that can learn from historical data, adapt to evolving network dynamics and make adequate decisions for improving the network performance and the energy efficiency [3]. The concrete specification of the supported AI/ML workflows in the O-RAN is still on-going [4]. Nevertheless, it will allow several options for providing AI/ML workflow services, for example model management, model training, model inference, data preparation, etc., at the Service Management and Orchestatration (SMO), the Non-RT RIC, the Near-RT RIC or through external components. Tightly (image-based) and loosely (file-based) coupled approaches will be also supported, allowing rApps/xApps to host the models and the training/inference runtimes or to use exposed AI/ML services provided by other components, respectively.

In this context, besides the user plane and data plane, BeGREEN introduces an Intelligent Plane, which allows the introduction of AI/ML control and management plane functions to reduce the overall energy consumption of the RAN infrastructure [5]. The proposed Intelligent Plane incorporates an AI Engine, which will provide a serverless execution environment hosting the AI/ML models, offering inference and training services to the rApps/xApps by following a loosely coupled approach.

The rest of the paper is organised as follows. Section II provides an overview of the related work in the context of open-source RIC implementations to implement automated control-loops. Section III presents the architecture of the BeGREEN Intelligent Plane and the AI Engine, with focus on the designed AI/ML workflows. Section IV presents the main components and workflows involved in a specific use case based on energy-efficient cell on/off control. Finally, conclusions are summarised in section V.

II. Related work

This section briefly presents relevant open-source RIC implementations and their utilisation to implement intelligent and automated control loops. The O-RAN Alliance and the Linux Foundation are collaborating by means of the O-RAN Software Community (OSC) to develop open-source Non-RT and Near-RT RIC solutions aligned with O-RAN specifications. In parallel, two additional initiatives, the Open Air