

EdgeRIC: Empowering Realtime Intelligent Optimization and Control in NextG Networks

Woo-Hyun Ko¹, Ushasi Ghosh², Ujwal Dinesha¹, Raini Wu²,

Srinivas Shakkottai¹ and Dinesh Bharadia²

¹ Texas A&M University, TX, USA, ² UC San Diego, CA, USA

{whko, ujwald36, sshakkot}@tamu.edu {ughosh, rainiwu, dineshb}@ucsd.edu

ABSTRACT

Radio Access Networks (RAN) are increasingly softwarized and accessible via data-collection and control interfaces. RAN intelligent control (RIC) is an approach to manage these interfaces at different timescales. In this paper, we develop a RIC platform called RICworld, consisting of (i) EdgeRIC, which is colocated, but decoupled from the RAN stack, and can access RAN and application-level information to execute AI-optimized and other policies in realtime (sub-millisecond) and (ii) DigitalTwin, a full-stack, trace-driven emulator for training AI-based policies offline. We demonstrate that realtime EdgeRIC operates as if embedded within the RAN stack and significantly outperforms a cloud-based near-realtime RIC (> 15 ms latency) in terms of attained throughput. We train AI-based policies on DigitalTwin, execute them on EdgeRIC, and show that these policies are robust to channel dynamics, and outperform queueing-model based policies by 5% to 25% on throughput and application-level benchmarks in a variety of mobile environments.

1 INTRODUCTION

NextG cellular networks must support applications ranging from interactive streaming media, AR/VR, control of robot systems, to industrial IoT 4.0. Many of these applications depend on tightly coupled chains of sensing, communication, and computation that requires wireless communication with low-latency, high-throughput, and high-reliability guarantees hitherto unavailable. The requirements are sometimes hard to quantify without contextual information about applications, such as nearness to stalling (media streaming), viewing angle (AR/VR), pose and location (robot control), age of information (industrial IoT). The problem is compounded by the fact that wireless links are fast-changing across diverse spectrum bands and user mobility patterns. To meet these requirements, traditional notions of tightly segregated layered networking are being revised in favor of highly adaptable cross-layer optimization of wireless resources.

NextG networks have the potential for autonomous adaptation to achieve such cross-layer optimization. With the

advent of O-RAN (Open Radio Access Network), NextG cellular networking is being powered by softwarization and disaggregation at all layers, enabling both the ability to run the RAN stack on a variety of distributed compute, and the means to monitor and control it via newly established interfaces as shown in Figure 1. In addition, the notion of RAN Intelligent Control (RIC) has arisen in parallel as an approach to managing these new interfaces. RICs are decoupled from the time-sensitive RAN stack, and the hope is for them to access both application layer and RAN-level information to provide cross-layer decision and control, as well as AI/ML capabilities to the RAN stack for supporting diverse and demanding applications.

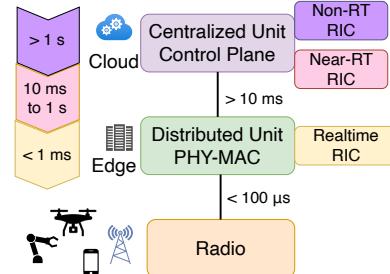


Figure 1: Timescales of RAN Intelligent Control for O-RAN. We desire realtime control at a latency < 1 ms.

The RAN-stack operates in realtime at a transmission time interval (TTI) of 1 ms or less (based on the 5G standards) i.e., it has to take control decisions at each ms. Recent attempts at RIC have been cloud-compute-based (i) near-realtime RIC (near-RT RIC), supporting a feedback loop to RAN stack at a timescale of 10 ms to 1 s, and (ii) non-realtime RIC (non-RT RIC), supporting a feedback loop at a timescale > 1 s. This cautious approach is taken to avoid impacting the time-sensitive tasks that must be completed within each TTI in the RAN stack. Breaching physical layer and medium access control (PHY-MAC) task-latency guarantees can lead to deleterious behaviors such as user equipment (UE) detachment. Hence, there is a mismatch between the realtime operations of the RAN and the current near-RT or non-RT RIC solutions.

The mismatch in RAN-level events and near-RT and non-RT RIC timescales implies that they can only undertake cross-layer decisions for prioritization of specific applications at a coarse level via resource slicing at the RAN. For instance, near-RT RIC may assign a streaming application to a resource-rich RAN slice, which might have the desired effect of reducing stalls because the RAN-level scheduler (which is oblivious to the application) provides low latency or high throughput using more resources available in that slice. However, such coarse slicing approaches could quickly become untenable with increasing loads being placed on wireless access networks by demanding applications, all of which cannot be assigned to such resource-rich slices.

Providing tight performance guarantees to demanding applications while maintaining resource efficiency implies that decision-making must be at the TTI timescale of underlying channel variations. Thus, truly transformative solutions for algorithmic control of PHY-MAC will require a *realtime* (TTI-scale) RIC located on edge compute near the radio head.

Main Contributions

In this paper, we propose RICworld, which can train and deploy policies for cross-layer information access and control at the sub-millisecond timescale. RICworld consists of two major components, namely, (i) a monitoring and control microservice called EdgeRIC, and (ii) an emulator called DigitalTwin. EdgeRIC is a microservice on edge compute, which is decoupled from the RAN stack and connects with it over an O-RAN like standard to execute AI-optimized and other policies in realtime (sub-millisecond). DigitalTwin is a cloud-compute based emulator, which enables the design, training and deployment of AI-optimized algorithms via full-stack emulation with multiple users and applications. RICworld can also leverage existing near-RT and non-RT RIC located on cloud-compute via an in-built information exchange.

Design Decisions

Where should EdgeRIC reside? An AI-optimized policy embedded within the RAN can access information in realtime, but also has a chance of breaking the latency-sensitive RAN stack, requires an application-aware RAN, and would require re-compiling the RAN for each algorithm update. On the other hand, decoupling completely from the RAN via cloud-based near-RT RIC has the issue of delayed information from the RAN over jittery, non-RT network links. Our key idea is that by co-locating with, but decoupling from the RAN via a realtime information exchange, EdgeRIC cannot impinge on latency-constrained RAN tasks, while yet accessing information and providing control in realtime. EdgeRIC hosts realtime applications called μ Apps with control policies that are agnostic to the RAN implementation. The design

is fault tolerant in that an absence of input from EdgeRIC means that the RAN continues with its usual operation.

What should be EdgeRIC’s action space? Many queueing systems have highly structured optimal scheduling policies that take the form of deciding the relative weight of one queue vs. another—a notion called indexing, with the Whittle index being the most well-known such policy class. In a likely first attempt in cellular networks, we leverage this observation to utilize a weight-based abstraction of control, under which each UE is assigned a weight by EdgeRIC at each TTI, which the RAN stack uses to determine their relative importance to perform resource allocation decisions. We then provide OpenAI Gym compatibility for μ Apps to execute standardized AI-optimized policies that output these weights.

How do we train cross-layer AI-optimized algorithms? AI algorithm training in a production system can result in excessive compute loads and extended training times. We equip RICworld with DigitalTwin that resides on cloud-compute at non-RT RIC, under which the core, RAN, RICs and multiple UE modules and their applications can be run over a trace-based emulation of channel quality index (CQI). AI/ML models are robustly trained under a variety of channel conditions, and are then sent for adaptation and realtime execution in μ Apps on EdgeRIC. Thus, μ Apps are able to update their policies in runtime as training and adaptation continue.

Performance Evaluation

We evaluate RICworld using the open source srsRAN stack by connecting to it via a custom variant of the O-RAN E2 application protocol (E2AP). We develop a custom scheduler placed within the srsRAN stack to perform proportional resource allocation using weights generated by EdgeRIC. We show that the RAN to EdgeRIC round trip time (RTT) has a median of only about $100\mu s$, and that scheduling algorithms operating at the TTI-timescale embedded within the RAN stack or as μ Apps show near-identical spectral efficiency, i.e., EdgeRIC operates as if from within the RAN stack and is capable of realtime decision and control.

Next, we develop AI-optimized μ Apps to provide resource allocation decisions each TTI. The challenge is to complete measurement, policy execution and resource allocation within the TTI. We train μ Apps via domain-randomized reinforcement learning (RL) over DigitalTwin to optimize performance of two types of applications, namely (i) throughput maximization, (ii) streaming media with stall minimization over a variety of channels. Training takes about ten minutes, while execution takes less than $150\mu s$. Hence, policy training and realtime feedback control are feasible on RICworld.

Finally, we conduct over-the-air experiments and trace-based emulations with devices on a turntable, mobile robots, cars and drones. Our baselines are well-established, queueing-model based algorithms, such as max-weight scheduling. We

first show that AI-optimized policies using the weight-based abstraction yield about 5% throughput gain over classical scheduling. We then show that in an environment where media streamers co-exist with other applications loading the system, RICworld can provide over 25% decrease in stalls. We conclude that AI-based realtime cross-layer optimization of cellular network resources is both achievable and valuable.

2 RELATED WORK

There has been significant recent work on developing near-RT RIC approaches, several of which use AI-optimized procedures for RAN control. For instance, Scope [4] provides a containerized approach for instantiating cellular network elements, an emulation module for testing in real-world scenarios, a data collection module for AI/ML applications, and APIs for control of network functionalities. The Colo-RAN [23] platform is a publicly available Open RAN solution that is based on the world's largest network emulator, Colosseum [15]. [5] conducted experiments on Colosseum to show the integration of a similar platform with deep RL agents. The focus of these works is on near-RT RIC based control, under which RAN resources are sliced in near-RT via an xApp and RAN-embedded schedulers are assigned to each one.

Other work explores the application space possible with near-RT RIC and non-RT RIC. In this context, [19] proposes an intelligent radio resource management scheme which leverages LSTM neural networks, non-RT and near-RT compute platforms, to learn the spatial pattern of data traffic and predict possible congestion. Based on these predictions, radio resources are dynamically re-allocated to prevent congestion and ensure optimal network performance. [32] presents a software-defined radio testbed featuring an open-source 5G system that interacts with the O-RAN near-RT RIC through standard interfaces with two xApps for RAN slicing.

Moving closer to the TTI timescale, ChARM [1] presents spectrum selection based on supervised learning over IQ samples utilizes data collection in realtime, but with control at near-RT RIC. For even finer control, [7] proposes an architecture for integrating distributed applications (dApps) into the O-RAN, and present simulations on the possible benefits of network intelligence at realtime (<10 ms). Further, [8] provides a conceptual messaging approach for enabling realtime RIC (<1 ms) and presents a feasibility study. In a similar framework, FlexRAN [9] provides a software-defined RAN platform, under which a master controller communicates with agents embedded in the RAN stack. The architecture requires agents tailored to the LTE stack, rather than functionally isolating them as O-RAN standards suggest. Although the master controller can host a MAC scheduler, FlexRAN does not show the ability to train or utilize AI-optimized policies while maintaining realtime constraints. FlexRIC [26]

is a more modular variant of FlexRAN. The focus is on simplifying the 5G near-RT RIC architecture, using the same agent-controller approach of FlexRAN.

In the applications domain, streaming media has received much attention for AI-optimized control. For instance, AI/ML for choosing video streaming rate selection is considered in [11, 14, 22, 33, 35] from the server's perspective, and assumes that the network provides packet delivery with certain statistics that can be learned. [3] studies optimal policies when the network can be controlled in the context of WiFi-based access. Here, non-RT reconfiguration of WiFi flow priorities using AI-optimized policies is shown to improve streaming performance. [20] investigates a system for streaming high-quality augmented reality (AR) content over wireless networks that experience rapid fluctuations in performance.

In contrast to the above work, ours is a simple, lightweight, decoupled approach towards ensuring that TTI-level synchronized policies can be trained in non-RT and executed in realtime. Specifically, we show that our approach provides the ability to train robust cross-layer optimized policies in non-RT and a guarantee of completing the full feedback loop from sensing, AI-based policy execution and control within each TTI, and are the first to verify our claims while running full stack over-the-air experiments on mobile nodes.

3 RICWORLD CONCEPT ARCHITECTURE

We develop a concept architecture for RICworld. Our goals for its constituents, EdgeRIC and DigitalTwin are as follows:

O-RAN Compatibility: EdgeRIC must be consistent with O-RAN messaging standards and RICworld must provide compatibility with near-RT and non-RT RIC. This will ensure ease of deployment and integration with existing RICs.

Realtime: EdgeRIC must be able to access RAN state information and provide control actions at each TTI. This will enable it to respond to a fast-changing wireless channel.

Robustness: EdgeRIC must not have any blocking elements that could slow down the RAN stack. This will ensure that any missing or delayed state or control does not break the tight constraints of TTI-scale RAN operations.

Reconfigurability: EdgeRIC must be able to update control policies in runtime. This will permit adaptation in response to data collected without having to restart the RAN or RIC.

Cross-layer Awareness: Apart from the RAN stack, EdgeRIC must be able to communicate with the applications running on UEs or other network elements. This will enable cross-layer optimization while accounting for application state.

Support for AI/ML Workflows: EdgeRIC must support standardized AI/ML codebases that might require offline,

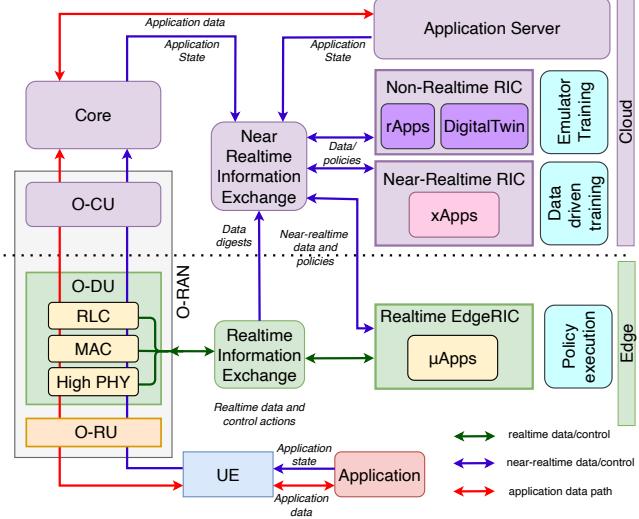


Figure 2: RICworld concept architecture, showing EdgeERIC, DigitalTwin and integration into O-RAN.

online, simulation or emulation-based data collection and training. This will ensure compatibility with a wide variety of AI/ML approaches for both inference and control.

Emulation Environment: DigitalTwin must operate over an identical stack of RAN, core, RIC(s), UEs and applications. It must accept data digests from over-the-air execution to ensure a realistic wireless and application environment. Queueing-model based or AI-optimized policies should be directly transferable to EdgeERIC without modification.

3.1 Realtime O-RAN Connectivity

EdgeERIC is illustrated in Figure 2, where we have shown it within the O-RAN architecture. O-RAN consists of a radio unit along with disaggregated microservices that perform the RAN functions. These microservices are divided across edge compute (near the radio) and cloud compute resources, based on the required latency targets. The components of the O-RAN stack are as follows: (i) RF Frontend: Open Radio Unit (O-RU), (ii) Edge Compute: Realtime components at the Open Distributed Unit (O-DU) supporting High-PHY, MAC and radio link control, and (iii) Cloud Compute: Open Centralized Unit (O-CU) with control and management functions. The final element is (iv) Cloud Compute: 5G Core, supporting management, billing and Internet gateway functions.

The O-RAN standard also provides specifications for two cloud-hosted microservices for (i) near-RT RIC, which supports xApps for policy adaptation at near-realtime (10 ms - 1 s), and (ii) non-RT RIC, which supports rApps providing microservices management and data analytics. The standard provides protocols for the near-RT and the non-RT RIC to communicate with the O-RAN stack and with each other. In particular, the E2 Application Protocol (E2AP) operates

over SCTP and provides pub-sub and on-demand messaging between RAN and near-RT RIC at near-RT latency.

We align our architecture with O-RAN, and conceptualize EdgeERIC as a microservice positioned at the O-DU for physical proximity to realtime RAN functionalities at the PHY-MAC level. EdgeERIC supports μ Apps representing realtime policy execution (TTI latency), and they can obtain RAN state information and apply control actions. EdgeERIC connects to the O-RAN stack via a realtime information exchange, which we visualize as a realtime-E2 protocol (RT-E2), which operates over inter-process communication (IPC) between microservices to ensure realtime latencies.

Our architecture also provides a near-RT information exchange, which can be used for data collection and sharing between RT, near-RT and non-RT RIC. This data can include policies that are developed at one of the RICs and is to be updated or executed at one of the others. The near-RT information broker can be visualized as a database that also provides the ability to store data digests for post processing.

3.2 Decoupled Execution, Runtime Updates and Cross-layer Optimization

The O-RAN stack must satisfy tight guarantees on PHY-MAC task completion times so as to maintain synchronization across all processes that have to be completed in each TTI. Breaching PHY-MAC latency guarantees can lead to unstable performance or deleterious behaviors such as UEs detaching. Although EdgeERIC operates in realtime, it is not integrated into the O-RAN PHY-MAC microservices at the O-DU. This allows us to decouple EdgeERIC operations with those of the O-DU stack. It can thus run on separate CPU cores from the O-DU microservices. The realtime information exchange ensures that the O-RAN stack never is in a blocking state waiting for input from EdgeERIC. Furthermore, events at EdgeERIC are driven by the TTI clock ticks from the O-DU, resulting in EdgeERIC being synchronized to the events at the O-DU stack.

Our decision to instantiate EdgeERIC as a decoupled realtime microservice at the O-DU enables it to support a variety of μ Apps that can be brought up and torn down at will. Each such μ App can utilize the realtime and near-RT information exchanges to obtain data to perform analytics, or to execute policies to generate control decisions. Since the O-DU stack is non-blocking on inputs from EdgeERIC, these μ Apps can be modified in runtime as data is collected and policies running on them need to be updated.

The decoupled μ Apps-based architecture is also compatible with cross-layer information sharing with EdgeERIC. While μ Apps utilize RT-E2 messages for communicating with the O-DU, they can also use other protocols in near-RT or non-RT for obtaining information about other elements of the

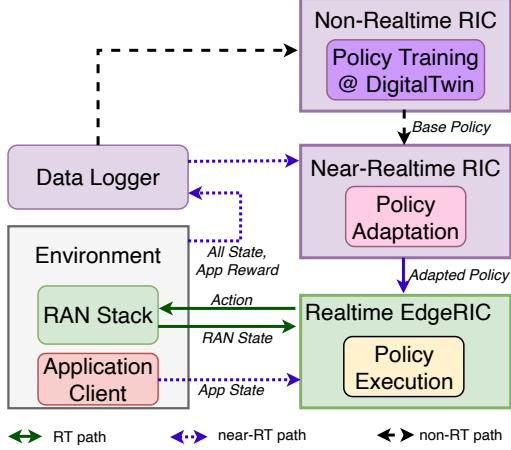


Figure 3: DigitalTwin based Non-RT policy training, Near-RT policy adaptation and RT policy execution.

applications stack. These protocols could include OpenFlow (network layer messaging), or a variety of application-level pub-sub protocols such as ROS (for robot sensing and control streams) or message queueing protocols (obtaining state information from AR/VR or media streaming applications). Hence, the EdgeRIC architecture can easily incorporate cross-layer policies that optimally control the PHY-MAC stack.

3.3 Emulator and AI/ML Workflow

EdgeRIC directly enables support for optimization based approaches to modulation, coding and queuing that are designed around well studied, substantiated, and tractable models. For instance, we can immediately instantiate approaches such as proportionally fair [28] or max-weight [31] scheduling across UEs on a per TTI basis with execution in RT, as if embedded within the RAN stack.

The architecture also supports approaches such as Reinforcement Learning (RL), a branch of ML that is explicitly tailored towards learning feedback-control policies. Here, the core idea is to “learn by doing” and to tailor the control policy in an online manner based on the information obtained thus far. The RL workflow is well aligned with the modality of a DigitalTwin based non-RT training of a base policy using data collected offline or via an emulator. Such policies can then undergo near-RT adaptation to the current environment, along with RT policy execution. Simultaneously, data is gathered at the edge, which is shared with the non and near-RT RIC for accurate training and adaptation. This three-timescale RL workflow is illustrated in Figure 3.

4 EDGERIC IMPLEMENTATION

We now describe the implementation of EdgeRIC to satisfy the goals of our concept architecture. The experimental results that we present in this section were collected on two

servers: Intel Xeon Gold 5218R CPU @ 2.10GHz, 20 cores and Intel i9 CPU @ 2.4GHz, 12 cores, without using GPUs.

4.1 O-RAN Compliant Messaging

O-RAN specifications provide the E2 interface between the near-RT RIC and the O-DU or O-CU. Specifically, the E2 application protocol (E2AP) operates over SCTP and provides near-RT services for RAN monitoring and control. E2 does not support a realtime connectivity service, and we extend the specification to create a realtime variant that we call RT-E2. RT-E2 supports two basic procedures to connect EdgeRIC with the PHY-MAC stack at the O-DU.

RT-E2 Report: This is a periodic pub-sub procedure under which a module at the O-DU, such as radio link control may publish information at a given rate. Our default periodicity is one TTI, i.e., information may be generated in realtime. μ Apps at EdgeRIC may subscribe to the RT-E2 Report service and utilize it for inference and control. Subscription may be blocking in that the μ App will proceed only when new information is available from the RAN.

RT-E2 Policy: This is an event-driven pub-sub procedure under which a μ App at EdgeRIC may publish information to one of the O-DU modules such as UE priorities for resource allocation at the MAC layer. This information is used directly for realtime control at the O-DU. Subscription is non-blocking in that the O-DU subscriber will move on if no new information is available on this procedure, without breaking the tight TTI deadlines required by PHY-MAC.

TTI-Level Sync Between RAN and EdgeRIC: It is critical for EdgeRIC to maintain TTI-by-TTI sync between RAN and RT-RIC in order that the control actions sent in realtime from the RT-RIC and the reward information (during training) from RAN accurately correspond to the current state at the RAN. Hence, we design sync procedures to ensure that RAN and RT-RIC maintain coherence. We maintain TTI counters at the RAN and RT-RIC, called *RANtime* and *RICtime*, which increase their counts at the completion of the tasks in that TTI at RAN or RT-RIC, respectively.

The two possibilities for asynchrony between RAN and RT-RIC and their respective solutions are: (i) **Lazy RAN**: Here, $RANtime < RICtime$, which can happen during system initialization. RT-RIC pauses execution until $RANtime = RICtime$ when a Lazy RAN event happens. (ii) **Lazy RIC**: Here, $RANtime > RICtime$, which can happen due to delayed computation of actions at the RT-RIC. The RT-RIC subscriber always keeps only the latest RAN message in its subscription queue, and sets $RICtime = RANtime$ when a Lazy RIC event happens. Note that the RAN must have a default option to deal with Lazy RIC events, since it might not receive any control input from the RT-RIC during these

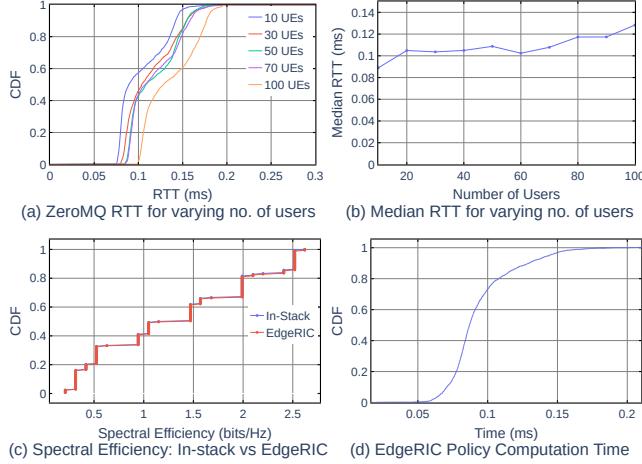


Figure 4: EdgeRIC Feedback latency, spectral efficiency and AI-policy execution times.

times. We will demonstrate that with the employment of such sync, the RAN and RT-RIC operate correctly in tandem, and feedback control and RL training proceed successfully.

4.2 Realtime Connectivity to RAN

We need a means of decoupling the RAN stack from EdgeRIC, and yet allow realtime information exchange between them. A message-passing layer can provide such decoupling to ensure that the strictly RT RAN stack is not negatively impacted by the introduction of additional RT components. We choose ZeroMQ [34] as the inter-process message passing framework over which we implement RT-E2 APIs for decoupled operation of between RAN and EdgeRIC. The low-latency and low-overhead nature of ZeroMQ is well suited for our architecture, which comprises of processes, both with strict RT requirements and looser near-RT requirements.

The fundamental timer in our system is the TTI-level clock tick from the RAN stack to which all processes must be synchronized. We therefore publish RAN-level information once every TTI, and use a RT-E2 Report type blocking subscription with the ZeroMQ conflate option set to ‘true’ (to keep only the latest RAN message in the queue) at RT-RIC to ensure synchronization of its actions at each TTI. All other subscriptions are set as non-blocking so that information is obtained whenever it is published, but no action is taken at RT-RIC until the clock tick at the RAN. Other modules, such as near-RT AI/ML training are self-clocking based on data digests received, with policy updates being published to the RT-RIC. Thus, we permit the coexistence of delay-constrained RT control loops and elastic data streams and policy updates in our architecture.

We conducted an evaluation of the suitability of ZeroMQ for RT operations by running srsRAN stack and the RT-RIC

on a server. The downlink channel operating at 10 MHz was fully loaded, with an over-the-air throughput between physical radios of approximately 37.5 Mbps during these experiments. The results are shown in Figure 4(a) and 4(b), which indicate that even while publishing information and receiving control actions for about a hundred UEs (physical and virtual), the median round trip latency of our implementation is an acceptable 100 μ s. In the context of resource block (RB) allocation at the MAC layer that we focus on as our sample application, this data acquisition path consists of the RAN stack publishing its state information on the active UEs, including their Radio Network Temporary Identifiers (RNTIs), Channel Quality Indicators (CQIs), backlog buffer states, their past downlink transmit bitrates, and RT-RIC responding with downlink RB allocation decisions for each UE. We see in Figure 4(c) that the spectral efficiency of running a policy in EdgeRIC vs. doing so with the same policy embedded within the RAN stack is essentially identical, i.e., the decoupled architecture is successful at maintaining efficiency without compromising on RAN stack robustness.

4.3 Cross-Layer Connectivity and Logging

Our choice of ZMQ for inter-process communication between RAN and EdgeRIC is also extendable to cross-layer application awareness. Since ZMQ is capable of also operating over TCP or UDP on an IP network, applications can simply use ZMQ to publish their state information to EdgeRIC. Apart from being lightweight and having APIs in most programming languages, ZMQ also permits client authentication and encryption via CurveZMQ [6]. This allows secure state-sharing with EdgeRIC.

We also enable RICworld with an in-memory Redis database for data logging and sharing. Redis is a fast, lightweight, key-value store, in which we log data digests, as well as trained models for sharing across the elements of RICworld. An added advantage of Redis is that we can save all traces to drive at experiment conclusion, which allows for post processing and performance analysis.

4.4 Weight-based Abstraction of Control

In resource constrained systems, such as wireless radio resource management, optimal policies often have simple and easy to learn structures, such as threshold policies [10, 16], index policies [21, 24, 25] and linear policies [2, 12]. In many other cases, the optimal value function may have similar simple and easy to learn structures, such as monotonicity or concavity [2]. [17] talks about assigning certain weights to prioritize flows, ensuring fairness among flows. All the above structured policies can effectively be represented by assigning relative priorities to the different connected UEs.

For example, the so-called Whittle index is a scalar parameter corresponding to the value of resources allocated to a given UE, which can quickly be learned independently of other UEs [18]. The Whittle indices of the UEs can then be used to prioritize resource block (RB) allocation to those UEs that have high indices. Resource allocation may also be done with a fairness metric in mind, such as proportional fairness, where RBs are assigned to a UE based on the ratio of its current channel quality as compared to its average channel quality, or max-min fairness [29]. Here too, the relative priority of a UE can be represented by a weight assigned to it.

Motivated by the above ideas, our general approach for RB allocation is for EdgeRIC to provide values w_i for each connected UE i , for both uplink and down link over realtime information exchange at each TTI. The 5G MAC will then allocate an number of RBs in a manner proportional to w_i over the next TTI. Such an abstraction provides simplicity of actions for the resource allocation policy, while maintaining its ability to attain near-optimal allocations in realtime.

4.5 Integration with OpenAIGym

OpenAIGym is an open source python library that provides a framework for developing an interface to interact with and query the environment by any given algorithm. While it is typically used for developing and comparing reinforcement algorithms, it can be used as a standard approach for realtime policy execution, regardless of whether the policy in question is based on reinforcement learning. This openness to the nature of the control algorithm motivates us to develop an OpenAIGym interface connecting the RAN stack, EdgeRIC and the control algorithms in the form of μ Apps that it hosts. This interface involves creating methods that communicate the current state of all the UEs, assign weights to UEs as suggested by the given policy, provide reward feedback for the current allocation, and reset the system when needed.

Developing the OpenAIGym interface allows for EdgeRIC to be seamlessly integrated with any optimization-based algorithmic approach, such as max-weight scheduling, or using a RL-optimized policy. Thus, the choice of OpenAIGym as our interface allows for swift policy development and freedom of execution of algorithms as desired. Figure 4(d) shows the time taken by EdgeRIC to execute a forward pass of a trained policy network using only CPU, while running a fully loaded RAN. The mean value is less than $100 \mu\text{s}$, which implies that TTI-scale execution is straightforward on general purpose compute.

5 DEPLOYING AI ON SRSRAN

We chose the open source Software Radio Systems srsRAN stack [30] as the experimental RAN system for EdgeRIC integration due to its simple, modular codebase, its stability

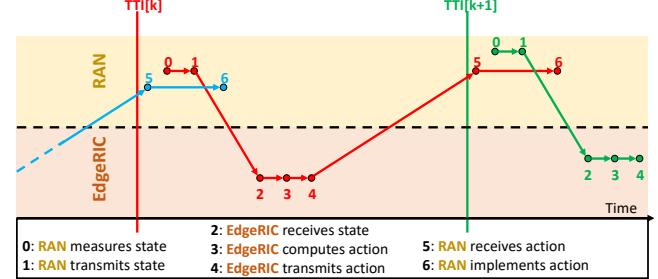


Figure 5: TTI-level events for EdgeRIC to RAN loop.

and compatibility with various core networks, 4G and 5G versions, and the availability of the srsUE codebase.

5.1 EdgeRIC and srsRAN

srsRAN runs on the general-purpose Ubuntu OS, which does not provide realtime guarantees. Its codebase is optimized to ensure that all fine-grain tasks are completed within the hard-constrained TTI boundaries. Thus, integrating of EdgeRIC with srsRAN requires caution to prevent disrupting time-sensitive event chains. Hence, we adopt a lightweight coupling approach prioritizing two main requirements;

RT-E2 API Support: We enable srsRAN codebase modules for radio link control and medium access control to utilize our custom RT-E2 APIs. These APIs run over ZeroMQ, which is already supported internally in the srsRAN codebase. Specifically, we use the RT-E2 Report API to collect Radio Network Temporary Identifiers (RNTIs), Channel Quality Indicators (CQIs), backlog buffer states, and their past down-link transmit bitrates for each UE, and publish these values each TTI for reception by EdgeRIC.

Weight-based RB Allocation: We create a custom scheduler for srsRAN that assigns RBs in proportion to weights provided for each UE. This scheduler is integrated with the RT-E2 Policy API, and subscribes to weights published by EdgeRIC every TTI in a non-blocking manner. It verifies correctness of the map between weights and RNTIs allocates RBs as prescribed. We maintain reliability by enabling it to use either a default scheduler or fixed weights, in case updates are not received from EdgeRIC in a particular TTI.

The full sequence of events is shown in Figure 5, where we see (i) state measurement and transmission from RAN, (ii) reception, processing and response at EdgeRIC, and (iii) final resource allocation at RAN. Note that srsRAN receives each EdgeRIC message well before the TTI boundary, but only reads it in a non-blocking manner at the beginning of each TTI. We show the CDF of end-to-end latency of the entire event chain from RAN to EdgeRIC and back (including policy execution via a forward pass on the policy network), culminating in resource allocation at each TTI in Figure 6.

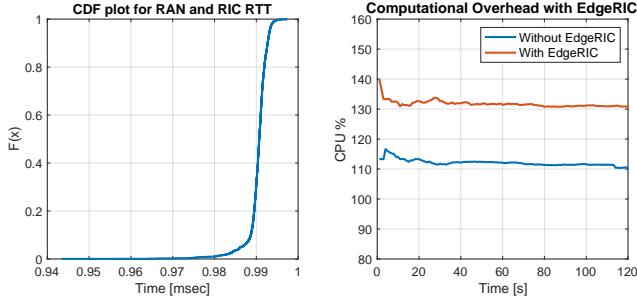


Figure 6: CDF of End-to-end RTT between RAN and EdgeRIC, showing that TTI timings are always met.

We observe that the end-to-end latency is always less than 1 ms, i.e., the RT-E2 sync procedure (Section 4.1) successfully meets the target of event completion within each TTI.

Finally, we measure the computational overhead of running EdgeRIC, as compared to running the vanilla srsRAN stack under a full traffic load. We see that the difference in CPU utilization is only about 20%, which means that EdgeRIC is fairly lightweight and does not need excessive additional compute resources. Hence, co-locating EdgeRIC at the O-DU level seems quite feasible.

5.2 Implementing DigitalTwin

A major challenge in RL-optimized control is the potential mismatch between the training environment and the real-world scenario, leading to a gap between the agent's training and real-world performance, known as the "sim-to-real" gap. Constructing a pure Python simulator to model the intricate relationships among the RAN, the RIC, and the end applications is a complex undertaking. Rather than building a simulator that models each of these components, we leverage virtual radios and channel simulation techniques to train RL policies directly on DigitalTwin which employs the real RAN and RIC codebases to replicate real-world relationships and performance accurately.

The construction of DigitalTwin is facilitated by existing support for ZeroMQ-based sinks and sources as virtualized radios in srsRAN. Complex-valued samples, typically sent to a software-defined radio for transmission, are instead sent to a specified ZeroMQ socket. Additionally, these samples can also be manipulated to introduce simulated channel effects to further bridge the sim-to-real gap of DigitalTwin. We define a GNU Radio flowgraph utilizing the ZMQ Source and ZMQ Sink blocks to multiplex the complex-valued samples to many srsUE inputs, routing them to the appropriate users. Finally, UEs and application servers can each have their own private IP namespaces, implying that real-world applications using TCP or UDP sockets can be run end-to-end on DigitalTwin, as outlined in Figure 7.

Table 1: RL Specifications: Throughput Maximization

State ($s[t]$)	$B_i[t], CQI_i[t] \forall i$
Action ($a[t]$)	$w_i[t] \forall i$
Reward ($r[t]$)	total throughput

Aside from the utilization of virtual radios and a simulated or trace-generated channel, all other components in DigitalTwin are clones of the ones used in a real deployment. Hence, DigitalTwin has very little sim-to-real gap.

5.3 Training RL Policies on DigitalTwin

We train an agent to perform optimal resource allocation using the model free RL algorithm, Proximal Policy Optimization (PPO) [27]. We chose PPO due to its ease of implementation and high sample efficiency. An iteration of training consists of collecting 5000 samples from the environment, adjustment of the agent's policy neural network weights through backpropagation, and the utilization of the updated agent to generate an additional 5000 samples to assess its performance and chart the training curve. A sample represents one transmission time interval (TTI) in the real system and consists of - the current state of the environment, action taken by the agent in this state, and the reward and next state observed as a result of this action. We may train RL policies on DigitalTwin for a variety of use cases, which can utilize either RAN information alone, or be augmented with application-level information for cross-layer optimization. Specially, we will focus on two use-cases on (i) downlink throughput maximization, discussed below and in Section 6.3, and (ii) video streaming stall minimization, discussed in Section 6.4.

With the a throughout-maximization objective, we choose the RAN-level state information as CQI and backlog buffer lengths for each UE, while the action is the allocation weights mentioned in Section 5.1, and the reward is the total throughput achieved. Reward saturation for this use-case was typically observed under 100 iterations of training, which equates to 500,000 training samples (TTIs). The RL framework specifications for each this setup is presented in Table 1. Here, $CQI_i[t]$ denotes the channel quality index of UE i at time t , $B_i[t]$ denotes the number of backlogged bytes in the downlink queue for UE i at time t , and the action returned is the weight $w_i[t]$ accorded to UE i at time t .

Including data collection/transfer overheads to the RL agent, and the time taken to perform actor-critic policy update at the end of each iteration, the total time taken for training completion in DigitalTwin is within approximately ten minutes. We present training results for this use-case on DigitalTwin below, while experiments on execution over EdgeRIC are shown in Section 6.3.

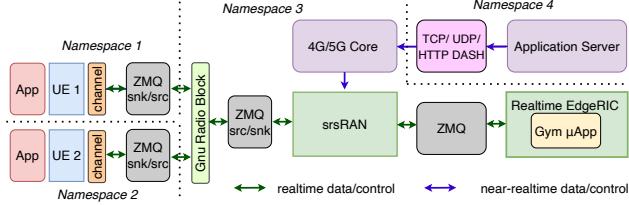


Figure 7: DigitalTwin emulation environment

5.4 Case Studies on DigitalTwin

We conducted emulations using synthetic channel traces in DigitalTwin to assess the potential gains achievable by a real-time agent for policy computation and control. Two metrics we use throughout our study is the downlink system throughput and the downlink backlog buffer lengths at the RAN. We desire to evaluate two basic questions, (i) *How much does performance improve with realtime control as opposed to near real time control* and (ii) *How does RL-based control policy perform compared to basic algorithms?*

We consider three basic algorithms for weight-based resource allocation. In all the below approaches, each UE is assigned a weight $w_i[t]$ at TTI t . The weights are then normalized over all UEs as $\tilde{w}_i[t] = w_i[t]/\sum_j w_j[t]$. The RAN receives the normalized weights $\tilde{w}_i[t]$ from the RIC, and performs an allocation of resource block groups (RBGs) in proportion to the weights, i.e., $R_i[t] = \tilde{w}_i[t]R_{total}[t]$, where $R_i[t]$ is the assignment to UE i , and $R_{total}[t]$ is the number of RBGs available in TTI t . While some approaches call for an absolute prioritization of UEs that have a maximum weight [31], we find in practice that a proportional division based on weight leads to better overall performances.

CQI-Fair Allocation: Here, the weight of UE is equal to its realized CQI. Hence, $w_i[t] = CQI_i[t]$, where $CQI_i[t]$ is the realized CQI of UE i at time t . This approach effectively tries to obtain a large total throughput by prioritizing these UEs that have a large CQI in the current timeslot.

Proportionally-Fair Allocation: Here, the weight of UE is the ratio between its current CQI and its average CQI, with the idea of prioritizing those UEs that have a good channel realization compared to their average value. The average CQI, denoted $AvgCQI_i[t]$ for UE i is calculated using an exponentially weighted moving average for each UE up to the current time, t . Thus, we have, $w_i[t] = CQI_i[t]/AvgCQI_i[t]$.

Max-weight Allocation: Here, the weight of a UE is the product of its current CQI and the backlogged bytes in the downlink queue corresponding to that UE. The max-weight policy is known to be throughput optimal [31], in that it can achieve the capacity region of the system. Thus, we have, $w_i[t] = CQI_i[t]B_i[t]$, where $B_i[t]$ is the number of backlogged bytes in the downlink queue of UE i .

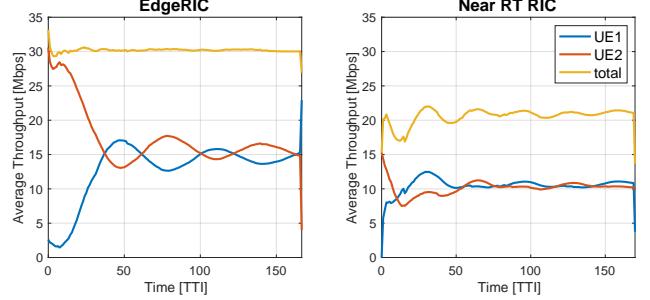


Figure 8: CQI-Fair Allocation with 2UE synthetic channels - one can support higher throughput while maintaining the stability of backlog buffers

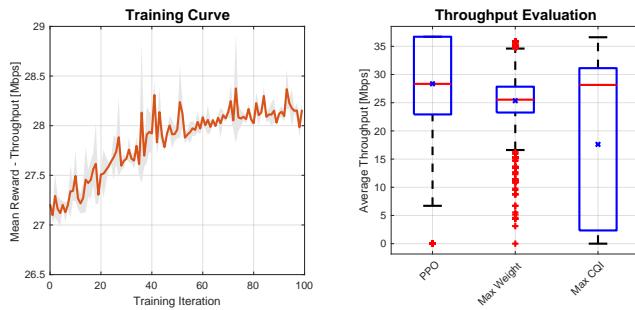
To answer the first question, the performance of downlink resource block (RB) allocation algorithms operating as μ Apps on EdgeRIC is compared to their operation as xApps on a cloud-based RIC, the distinction being the latency in state capture and control. A μ App at the edge experiences a low round-trip latency of a few tens of microseconds between receiving state information from the RAN and action generation, whereas an xApp in the cloud experiences both forward and reverse network latency between the RAN and the cloud, resulting in a round-trip time of tens of milliseconds. To demonstrate the effect of cloud latency, appropriate delays are artificially induced in DigitalTwin.

We begin our experiment with synthetic channel traces. The results of the CQI-Fair allocation algorithm, when executed as a μ App over EdgeRIC or as a Cloud-based RIC with a total monitoring and control latency of 30ms, are depicted in Figure 8. The findings suggest that the utilization of EdgeRIC leads to a 50% increase in throughput while preserving the stability of backlog buffers. This superiority in throughput performance is further illustrated in Figure 13 in the appendix, which portrays the results of a 4-user scenario. The results are summarized in Tables 2, 8. We see that real-time control significantly improves performance metrics in most cases.

To answer second question, we demonstrate the training and evaluation of an RL algorithm on the DigitalTwin. Scenario 1 is based on a 2UE environment. both connected users have uniform variation in CQI values over time, ranging from 1 to 15. In scenario 2, one user experiences good channel conditions (CQI values between 8 and 15), while the other user experiences poor channel conditions (CQI values between 1 and 7). In Scenario 3, the CQI values are randomly generated. Figure 9 shows training and evaluation on Scenario 2 and Table 9 summarizes the performance of RL algorithms on synthetic channel traces.

Table 2: Load: 35Mbps, Channel: 2 UE synthetic channel

EdgeRIC 15ms 30ms				
Max CQI	Avg. Thrpt.	32.6	24.2	18.0
BL[MB]	0.61	0.64	0.57	
Prop. Fair.	Avg. Thrpt.	30.7	25.7	21.9
	BL[MB]	0.65	0.67	0.68
Max Weight	Avg. Thrpt.	30.0	23.3	20.9
	BL[MB]	0.60	0.62	0.65

**Figure 9: RL training and evaluation on a synthetic channel trace**

6 RICWORLD EVALUATION

In this section, we evaluate EdgeRIC by addressing three fundamental questions. These questions are: (i) Does real-time RAN control through μ Apps on EdgeRIC outperform near-real-time control through a cloud-based RIC, (ii) Is it feasible to implement real-world AI optimization (RL training and feedback) for resource allocation over EdgeRIC, and (iii) Does application state feedback to EdgeRIC provide significant improvement to the end-user Quality of Experience (QoE)?

To evaluate the above questions, we first present our base-station setup and four types of UEs, from fixed to highly mobile users. Next, we discuss a variety of experiments conducted to result in dynamic wireless channels. We used these channel traces to evaluate the performance of traditional schedulers and an RL-based scheduler run as μ Apps on EdgeRIC. We also present results from real-world over the air experiments to support our claims. Finally, we provide a case study of a streaming application that benefits from this platform. Our RL-based scheduling algorithm outperformed traditional methods, confirming EdgeRIC's capabilities.

6.1 Experimental Setup

The EdgeRIC implementation is a modular extension to the srsRAN codebase, which can accommodate both software-defined radios and off-the-shelf UEs. Our base-station stack

was built using srsRAN version 21.10 to support EdgeRIC, and we utilized USRP B210 SDRs, an edge DU, and an embedded delay of 20ms in the CU stack. Our srsRAN base-station includes a logging mode for channel trace logging. To avoid interference and to move freely without disrupting building WiFi or commercial cellular transmissions, our experiments were conducted within the 2.5 GHz EBS band under experimental FCC license. Our typical experiments were with 10 MHz bandwidth, with a fixed base station and mobile users.

6.1.1 User Devices and Channel Traces. We use a static base-station with different UE types to collect channel traces as shown in Fig. 10. We considered various mobility models, both within the laboratory and in outdoor environments. We extract the channel quality indicator (CQI) values sampled at each transmission time interval (TTI) for each user at the srsRAN for runs of approximately 8 minutes each. Next, we summarize the UEs considered in this study.

TurnTable UE: Figure 10(a) shows the setup of a UE on a movable turntable, similar to a user sitting on a chair and rotating. The base station equipped with a B210 is on a static table, while a UE node with a B210 is mounted on the turntable. The UE node is moved away from or towards the base station at about 1 m/s and rotated to produce significant variations in its CQI values. The distance between the UE node and the base station ranges from 0.5 meter to 4 meter.

Car UE: Figure 10(b) shows a UE setup in a car, using a USRP B210 and laptop powered by the car. CQI data was collected while driving along three paths: 1) a 6-meter-radius circle with a maximum speed of 4.5 m/s, 2) a 30 meter x 3 meter rectangle with a maximum speed of 5.4 m/s, and 3) a 50-meter straight line with a maximum speed of 9 m/s.

Drone UE: We used a drone experiment to demonstrate performance with faster channel variations in 3D space. Figure 10(c) shows a B210 and a NUC (a small and portable computer) mounted on a Big-Hexy hexa-copter drone, while the base station, equipped with another B210, was placed on the ground. The drone was flown over the base station along two paths: 1) a 10-meter straight line at a height of 15 meters and with a speed of 2.2 m/s, and 2) a 15-meter straight line at a height of 20 meters and with a speed of 3 m/s.

Indoor robotic UE: Mobile robots were used to for indoor mobility experiments. In Figure 10(d), a B210 and laptop were mounted on a Jackal robot, while the base station with an X310 was placed on a table. The robot moved along a 1.6 m x 1.6 m square path and rotated at each corner, causing CQI values to drop due to signal blockage. The working area was limited to 4 m from the base station, similar to robot control or industrial IoT with Private 5G.

We summarize the results in Figure 10(e), which shows the CDF of the time interval between CQI changes. We see that several mobility scenarios yield CQI changes in the sub

Table 3: Summary of all scenarios

Scenario	Channel Description
Channel Traces from Experiments	
Scenario 1	2 Drone UEs
Scenario 2	2 Turntable UEs
Scenario 3	2 Car UEs and 2 Drone UEs
Scenario 4	2 Car UEs and 2 Indoor Robotic UEs
Scenario 5	2 Random Walk UEs and 2 Turntable UEs
Complete Over-the-Air Experiments	
Scenario 6	2 UEs on indoor mobile robots
Scenario 7	2 UEs on indoor stationary robots

Table 4: Load: 35Mbps, Channel Trace: 4 Turntable UEs

EdgeRIC 50ms 100ms				
Max CQI	Avg. Thrpt.	33.4	21.2	29.5
	BL[MB]	1.34	0.84	1.12
Prop. Fair.	Avg. Thrpt.	28.6	26.6	23.5
	BL[MB]	1.20	1.29	0.93
Max Weight	Avg. Thrpt.	33.2	28.8	31.0
	BL[MB]	1.14	1.30	1.12

10 ms range, with the drone trace showing this effect for almost 90% of samples.

6.1.2 Evaluation Scenarios. To generate realistic scenarios, we collected channel traces from various environments and utilized them in our DigitalTwin emulation setup. This setup includes the same core, radio access network (RAN), and user equipment (UE) modules as our over-the-air experiments that generated the CQI traces. By replaying the CQI traces with a desired number of UEs, algorithmic methods can be compared while maintaining the same end-to-end applications and channel conditions. For end-to-end over-the-air experiments, we used a X310 as a base station and two B210s as UEs. One of the UEs had lower channel quality than the other by placing them at different distances from the base station. Table 3 summarizes the various scenarios considered in this study.

6.2 Micro-benchmarks: Edge vs. Cloud

We utilize iPerf for measuring network throughput and generate microbenchmarks. We test our scenarios on different UEs, each with varying traffic loads, and report the throughput and backlog buffers observed with different scheduling algorithms presented in Section 5.4.

The validity of our results presented in Section 5.4 is confirmed using realistic channel traces (4 Turntable UEs), as

depicted in Table 4 Figure 11(a). This provides robust evidence in support of the hypothesis that real-time control can significantly improve the system throughput.

6.3 Impact of RL on Micro-benchmarks

We present compelling evidence for the feasibility, impact, and robustness of RL in executing real-time control policies. We begin by addressing the question of RL training in a digital twin environment, which is answered by the set of curves in Figure 15 of the appendix. These curves demonstrate that RL training with realistic channel traces is highly efficient, with an agent typically converging in just 20 to 40 iterations. The potential of RL is evident from a series of figures that answer important questions about its application.

Figure 11(b) shows us that RL can achieve higher throughput than traditional algorithms, bringing us to answer: is RL truly useful? This is further summarized in Table 5, which displays the total system throughput and the mean of the total backlog buffer for various scenarios, offering a snapshot of the RL performance in real-time. The values displayed in the table represent the throughput in Mbps (left) and the backlog buffer in MBytes (right).

Figure 11(c) offers an answer to the question of model transferability. It demonstrates that an RL model robustly trained in one scenario can be effectively used in another, revealing its generalizability to other scenarios with similar numbers of connected users. A model trained on the random walk scenario was used to evaluate on a 2 Drone UEs scenario (Scenario 1) and on a 2 Car UEs and 2 Turntable UEs scenario (Scenario 2). Yet, the third question is arguably the most important: can RL really provide performance gains with realtime wireless control in the real-world? Figure 11(d) offers a clearly affirmative answer. While conducting real-world experiments with EdgeRIC, we observed that excessive channel variations led to unreliable performance of traditional algorithms like max weight and max cqi, despite their theoretical effectiveness. This resulted in limited data transfer. However, the robustness of the RL policy ensured system stability and delivered a throughput that matched the load.

6.4 Cross-Layer Optimization: Case Study

We next consider the potential for enhancing application performance via cross-layer optimizations in a video streaming case study. To this end, we designed a heterogeneous application environment comprising four users. Two are linked to a HTTP video server, utilizing a high-quality video streaming application with a constant bit rate, while the others act as system loaders, simulating a background file downloader via iPerf traffic with 10 Mbps system load.

Our study uses the open-source GPAC [13] library as the video streaming platform, generating 2-second DASH-style

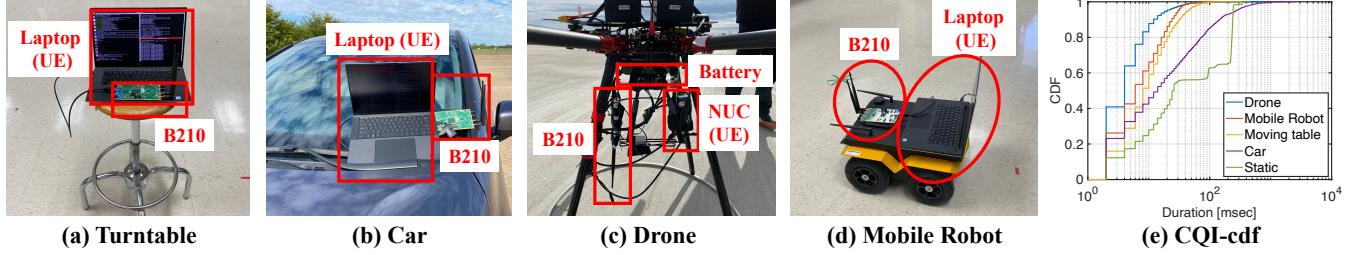


Figure 10: Experimental setups for collecting CQI data of various mobility and CDFs of their CQI traces

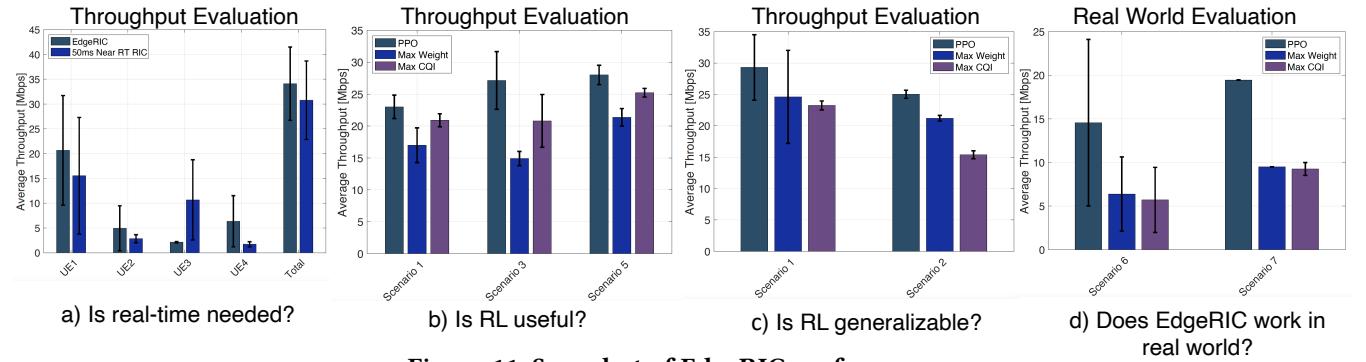


Figure 11: Snapshot of EdgeRIC performance

Table 5: Throughput and Backlog Buffer Evaluation

	PPO	Max Weight	Max CQI
Realistic Channel Traces			
Scenario 1	29.1/0.38	26.1/0.53	14.9/0.39
Scenario 2	30.5/0.38	31.9/0.43	14.42/0.39
Scenario 3	25.3/1.5	22.9/1.3	18.67/ 0.97
Scenario 4	25.9/1.5	23.9/1.21	20.3/ 1.05
Scenario 5	28.5/0.96	26.3/1.46	23.3/1.01
Over the Air Experiments			
Scenario 6	14.6/0.19	6.4/0.45	5.7/0.44
Scenario 7	19.33/0.05	10.71/0.34	9.06/0.35

segments at the video server, with a media buffer size of 6 seconds at the client end. The video frames display at a rate of 24 frames per second. When the media buffer falls below 2 seconds, it is considered a stall event. The stall duration is computed until the buffer refills to over 2 seconds and is ready for playback again. We updated the media buffer size information at the client every time a frame is about to start, every 40 milliseconds in a 24fps video, and published it to EdgeRIC. This allowed the EdgeRIC learning agent to be informed of the application state in the form of the media buffer size every 40 ms, along with network states at each TTI. To emulate a realistic scenario, we introduced a delay

Table 6: RL Specifications: Video Streaming

State ($s[t]$)	$B_i[t], CQI_i[t], MB_i[t] \forall i$
Action ($a[t]$)	$w_i[t] \forall i$
Reward ($\sum_i r_i[t]$)	$r_i[t] = \begin{cases} -20, & \text{if } MB_i[t] < 2 \text{ sec} \\ +2, & \text{otherwise} \end{cases} \forall i$

of around 20ms to simulate the uplink latency on srsRAN, as the media buffer size information needs to traverse through the core to the server before it becomes available to EdgeRIC.

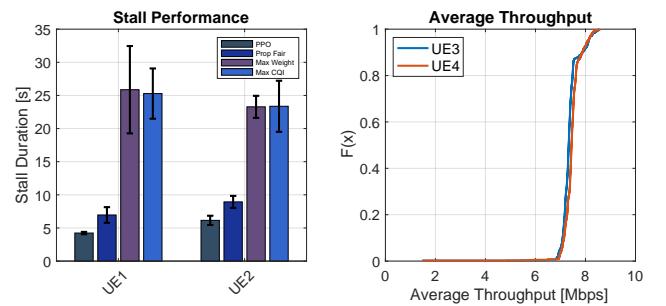


Figure 12: Application Performance

An RL agent is trained on DigitalTwin to execute realtime policies for a video streaming application in a heterogeneous environment with EdgeRIC, guided by a reward function

Table 7: Total Stall Duration: Video Streaming with EdgeRIC in real world experiments

PPO Prop.	Fairness	Max Weight
Run 1	2.6s	4.2s
Run 2	2.3s	3.8s
		4.5s

that aims to achieve optimal video playout without stalling and efficient resource allocation to system loaders. The RL framework specifications for this setup are presented in Table 6. Here, apart from downlink backlog $B_i[t]$, and channel quality $CQI_i[t]$ of UE i , we augment the state with the length of the media buffer (in seconds) of the video streaming application of UE i , denoted $MB_i[t]$. The reward now depends on the stall performance of the applications, with smooth playout receiving a positive reward, and a stall receiving a large negative reward. We found after the training process that the RL agent rapidly converged to an optimal policy within ten iterations. The RL policy effectively allocated sufficient resources to UE3 and UE4, with an average throughput of about 7.5 Mbps per user, maintaining a stable backlog buffer.

We next conduct real-world experiments on the video streaming application. Comparing the video streaming performance of UE1 and UE2 controlled by the RL policy and the standard RAN scheduling algorithms in real-time, the RL policy outperformed the standard algorithms by incorporating "application awareness," demonstrating its potential to provide a quality of experience (QoE)-optimized solution with proper training. Figure 12 summarizes this statement showing the metrics observed by playing a video for 120s. Table 7 presents the results of our experiment, which evaluated the stall performance of the video streaming application over the air in a static environment with a one-streamer and one-loader scenario for 30 s videos. We see that the RL trained policy only has about a third of the stalls experienced by the other approaches.

7 LIMITATIONS AND FUTURE WORK

We presented a platform RICworld, with a realtime controller microservice, EdgeRIC and a full stack emulator, DigitalTwin, which, when combined together, can train deploy cross-layer AI-optimization algorithms that can provide decision and control at a millisecond scale.

Extension across different ORAN stacks: Our work has demonstrated that simple modification to srsRAN (open source stack) enables us to support EdgeRIC. Going forward, will open-source our implementation for the community to re-create the results and scale it to different black box and white box implementations of the RAN stack for deployment.

Extension to different applications: We are probably the first to show end-to-end realtime over-the-air experiments

that demonstrate the feasibility and value of RL-based policies operating over a 5G stack. We do acknowledge our system needs testing with multiple applications, and open-sourcing the platform will help community efforts to do so.

Scale to 100s of users: Our demonstrations are with four users, primarily due to the instability with our setup of srsRAN for a larger number of UEs. However, we have shown that EdgeRIC platform can send the messages for 100 UEs, with a RAN to EdgeRIC latency of about $100\mu\text{s}$ (Figure 4). That said, we have not tested over-the-air with support for many more simultaneous UEs, which is key future work.

Mutliple base-stations: Our work has demonstrated performance improvements for a single base-station. We anticipate with multiple base-stations, we would have an EdgeRIC attached to each base-station/DU, with Near-RT RIC coordinating all the EdgeRIC instantiations. This would involve problems in federated learning that we will explore.

Ethical concerns: This work does not raise any ethical issues.

REFERENCES

- [1] Luca Baldesi, Francesco Restuccia, and Tommaso Melodia. 2022. ChARM: NextG Spectrum Sharing Through Data-Driven Real-Time O-RAN Dynamic Control. *arXiv:2201.06326* (2022).
- [2] Dimitri P Bertsekas. 2017. *Dynamic programming and optimal control*. Vol. 1,2. Athena scientific Belmont, MA.
- [3] Rajarshi Bhattacharyya, Archana Bura, Desik Rengarajan, Mason Rumuly, Bainan Xia, Srinivas Shakkottai, Dileep Kalathil, Ricky KP Mok, and Amogh Dhamdhere. 2021. QFlow: A Learning Approach to High QoE Video Streaming at the Wireless Edge. *IEEE/ACM Transactions on Networking* 30, 1 (2021), 32–46.
- [4] Leonardo Bonati, Salvatore D’Oro, Stefano Basagni, and Tommaso Melodia. 2021. SCOPE: an open and softwarized prototyping platform for NextG systems. In *Proceedings of the 19th Annual International Conference on Mobile Systems, Applications, and Services*. 415–426.
- [5] Leonardo Bonati, Salvatore D’Oro, Michele Polese, Stefano Basagni, and Tommaso Melodia. 2021. Intelligence and Learning in O-RAN for Data-Driven NextG Cellular Networks. *IEEE Communications Magazine* 59, 10 (2021), 21–27. <https://doi.org/10.1109/MCOM.101.2001120>
- [6] CurveZMQ 2023. CurveZMQ. <http://curvezmq.org/>. (2023).
- [7] Salvatore D’Oro, Michele Polese, Leonardo Bonati, Hai Cheng, and Tommaso Melodia. 2022. dApps: Distributed Applications for Real-Time Inference and Control in O-RAN. *IEEE Communications Magazine* 60, 11 (2022), 52–58. <https://doi.org/10.1109/MCOM.002.2200079>
- [8] Anonymous et. 2022. Anonymous to preserve double blind review. (2022).
- [9] Xenofon Foukas, Navid Nikaein, Mohamed M. Kassem, Mahesh K. Marina, and Kimon Kontovasilis. 2016. FlexRAN: A Flexible and Programmable Platform for Software-Defined Radio Access Networks. In *Proceedings of the 12th International on Conference on Emerging Networking EXperiments and Technologies (CoNEXT ’16)*. Association for Computing Machinery, New York, NY, USA, 427–441. <https://doi.org/10.1145/2999572.2999599>
- [10] Yu-Pin Hsu, Navid Abedini, Natarajan Gautam, Alex Sprintson, and Srinivas Shakkottai. 2014. Opportunities for network coding: To wait or not to wait. *IEEE/ACM Transactions on Networking* 23, 6 (2014), 1876–1889.

- [11] Tianchi Huang, Rui-Xiao Zhang, Chao Zhou, and Lifeng Sun. 2018. QARC: Video quality aware rate control for real-time video streaming based on deep reinforcement learning. In *Proceedings of the 26th ACM international conference on Multimedia*. 1208–1216.
- [12] Panqanamala Ramana Kumar and Pravin Varaiya. 2015. *Stochastic systems: Estimation, identification, and adaptive control*. Vol. 75. SIAM.
- [13] Jean Le Feuvre, Cyril Concolato, and Jean-Claude Moissinac. 2007. GPAC: Open Source Multimedia Framework. In *Proceedings of the 15th ACM International Conference on Multimedia (MM '07)*. Association for Computing Machinery, New York, NY, USA, 1009–1012. <https://doi.org/10.1145/1291233.1291452>
- [14] Hongzi Mao, Ravi Netravali, and Mohammad Alizadeh. 2017. Neural adaptive video streaming with pensieve. In *Proceedings of the Conference of the ACM Special Interest Group on Data Communication*. ACM, 197–210.
- [15] Tommaso Melodia, Stefano Basagni, Kaushik R. Chowdhury, Abhimanyu Gosain, Michele Polese, Pedram Johari, and Leonardo Bonati. 2021. Colosseum, the World’s Largest Wireless Network Emulator. In *Proceedings of the 27th Annual International Conference on Mobile Computing and Networking (MobiCom ’21)*. Association for Computing Machinery, New York, NY, USA, 860–861. <https://doi.org/10.1145/3447993.3488032>
- [16] Arupa Mohapatra, Natarajan Gautam, Srinivas Shakkottai, and Alex Sprintson. 2014. Network coding decisions for wireless transmissions with delay consideration. *IEEE Transactions on Communications* 62, 8 (2014), 2965–2976.
- [17] Kanthi Nagaraj, Dinesh Bharadia, Hongzi Mao, Sandeep Chinchali, Mohammad Alizadeh, and Sachin Katti. 2016. NUMFabric: Fast and Flexible Bandwidth Allocation in Datacenters. In *Proceedings of the 2016 ACM SIGCOMM Conference (SIGCOMM ’16)*. Association for Computing Machinery, New York, NY, USA, 188–201. <https://doi.org/10.1145/2934872.2934890>
- [18] Khaled Nakle, Santosh Ganji, Ping-Chun Hsieh, I-Hong Hou, and Srinivas Shakkottai. 2021. NeurWIN: Neural Whittle Index Network For Restless Bandits Via Deep RL. In *Thirty-Fifth Conference on Neural Information Processing Systems*.
- [19] Solmaz Niknam, Abhishek Roy, Harpreet S. Dhillon, Sukhdeep Singh, Rahul Banerji, Jeffery H. Reed, Navrati Saxena, and Seungil Yoon. 2020. Intelligent O-RAN for Beyond 5G and 6G Wireless Networks. (2020). <https://doi.org/10.48550/ARXIV.2005.08374>
- [20] Hyunmin Noh and Hwangjun Song. 2021. HA2RS: HTTP Adaptive Augmented Reality Streaming System. *IEEE Transactions on Mobile Computing* (2021), 1–1. <https://doi.org/10.1109/TMC.2021.3132665>
- [21] Ali ParandehGheibi, Muriel Médard, Asuman Ozdaglar, and Srinivas Shakkottai. 2010. Access-network association policies for media streaming in heterogeneous environments. In *49th IEEE Conference on Decision and Control (CDC)*. IEEE, 960–965.
- [22] Jounsup Park, Philip A. Chou, and Jenq-Neng Hwang. 2019. Rate-Utility Optimized Streaming of Volumetric Media for Augmented Reality. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems* 9, 1 (2019), 149–162. <https://doi.org/10.1109/JETCAS.2019.2898622>
- [23] Michele Polese, Leonardo Bonati, Salvatore D’Oro, Stefano Basagni, and Tommaso Melodia. 2022. ColO-RAN: Developing Machine Learning-based xApps for Open RAN Closed-loop Control on Programmable Experimental Platforms. *IEEE Transactions on Mobile Computing* (July 2022), 1–14.
- [24] Vivek Raghunathan, Vivek Borkar, Min Cao, and P Roshan Kumar. 2008. Index policies for real-time multicast scheduling for wireless broadcast systems. In *IEEE INFOCOM 2008-The 27th Conference on Computer Communications*. IEEE, 1570–1578.
- [25] Javad Razavilar, KJ Ray Liu, and Steven I Marcus. 2002. Jointly optimized bit-rate/delay control policy for wireless packet networks with fading channels. *IEEE Transactions on Communications* 50, 3 (2002), 484–494.
- [26] Robert Schmidt, Mikel Irazabal, and Navid Nikaein. 2021. FlexRIC: An SDK for next-Generation SD-RANs. In *Proceedings of the 17th International Conference on Emerging Networking EXperiments and Technologies (CoNEXT ’21)*. Association for Computing Machinery, New York, NY, USA, 411–425. <https://doi.org/10.1145/3485983.3494870>
- [27] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. 2017. Proximal Policy Optimization Algorithms. (2017). <https://doi.org/10.48550/ARXIV.1707.06347>
- [28] Stefania Sesia, Issam Toufik, and Matthew Baker. 2011. *LTE-the UMTS long term evolution: from theory to practice*. John Wiley & Sons.
- [29] S. Shakkottai and R. Srikanth. 2007. Network optimization and control. *Foundations and Trends in Networking* 2, 3 (2007), 271–379.
- [30] srsRAN, Inc. 2023. srsRAN, Inc. <https://www.srlte.com/>. (2023).
- [31] Leandros Tassiulas and Anthony Ephremides. 1990. Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks. In *29th IEEE Conference on Decision and Control*. IEEE, 2130–2132.
- [32] Pratheek S. Upadhyaya, Aly S. Abdalla, Vuk Marojevic, Jeffrey H. Reed, and Vijay K. Shah. 2022. Prototyping Next-Generation O-RAN Research Testbeds with SDRs. (2022). <https://doi.org/10.48550/ARXIV.2205.13178>
- [33] Gongwei Xiao, Muhong Wu, Qian Shi, Zhi Zhou, and Xu Chen. 2019. DeepVR: Deep Reinforcement Learning for Predictive Panoramic Video Streaming. *IEEE Transactions on Cognitive Communications and Networking* 5, 4 (2019), 1167–1177.
- [34] ZeroMQ 2023. ZeroMQ. <https://zeromq.org/>. (2023).
- [35] Yuanxing Zhang, Pengyu Zhao, Kaigui Bian, Yunxin Liu, Lingyang Song, and Xiaoming Li. 2019. DRL360: 360-degree Video Streaming with Deep Reinforcement Learning. In *IEEE INFOCOM 2019-IEEE Conference on Computer Communications*. IEEE, 1252–1260.

A APPENDIX

In the appendix, we provide additional results for different scenarios, which are not critical to the performance but provide visibility into the system.

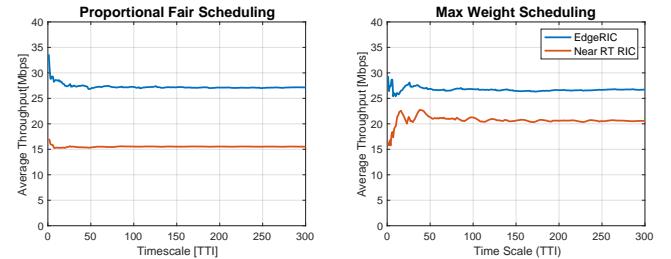


Figure 13: EdgeRIC Performance in a 4UE synthetic channel scenario.

Figure 13 shows that EdgeRIC outperforms Near Real-time RIC for implementation of microbenchmarks. This is the scenario with 4UEs on a synthetic channel trace. We implemented Proportional Fair Scheduling and Max Weight Scheduling as a μ App in EdgeRIC which exchanges state information and actions with RAN Stack over ZeroMQ-based communication. In order for performance comparison, we

implemented Near RT RIC by imposing delays in message deliveries between RIC and RAN. We used synthetic CQI traces for 4 UEs and evaluated each algorithm's performance by comparing their average throughput. The graph on the left side shows average throughput of Proportional Fair Scheduling and the graph on the right side shows one of Max Weight Scheduling. A blue line in the graphs are for EdgeRIC while a red line for Near RT RIC. As Near RT RIC showed low average throughput due to the delay, EdgeRIC successfully supported those microbenchmarks to achieve their best throughput.

Table 8: Load: 30Mbps, Channel: 4UE synthetic channel

EdgeRIC 15ms 30ms				
Max CQI	Avg. Thrpt.	26.3	16.4	15.9
BL[MB]		1.22	1.25	1.27
Prop. Fair.	Avg. Thrpt.	24.27	22.72	20.15
BL[MB]		1.37	1.08	1.27
Max Weight	Avg. Thrpt.	25.4	19.8	19.1
BL[MB]		1.33	1.05	0.99

Table 9: Throughput and Backlog Buffer evaluation of synthetic traces

PPO	Max Weight	Max CQI
Synthetic Channel Traces		
Scenario 1 27.8/0.38	25.6/0.38	19.0/0.38
Scenario 2 27.5/0.33	24.7/0.51	18.9/0.38
Scenario 3 26.8/0.38	25.2/0.51	25.2/0.76

Figure 14 shows the real-time CQI traces we collected to characterize various CQI mobility in different environments. An x-axis is time in TTI unit and a y-axis is CQI showing real-time channel quality. The report period of CQI values was 2 TTI period which is approximately 2 mili-seconds. While a mobile robot has some drops in its CQI traces due to the radio block by the lid of the laptop mounted on the mobile robot when rotating, drone's swift motions caused radical ups and downs in its CQI traces. In car's CQI traces, its variation has a characteristic of a long period and smooth curves because of its slow acceleration and deceleration. For a turntable, we was able to generate fast angular acceleration which caused drastic and kind of periodic changes in its CQI traces by fast rotation.

Figure 15 shows the training curves of RL PPO policy model and its throughput evaluations on DigitalTwin. The

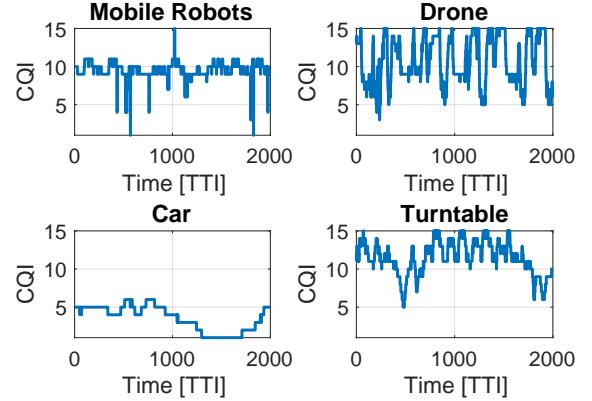


Figure 14: CQI Trace for different UEs

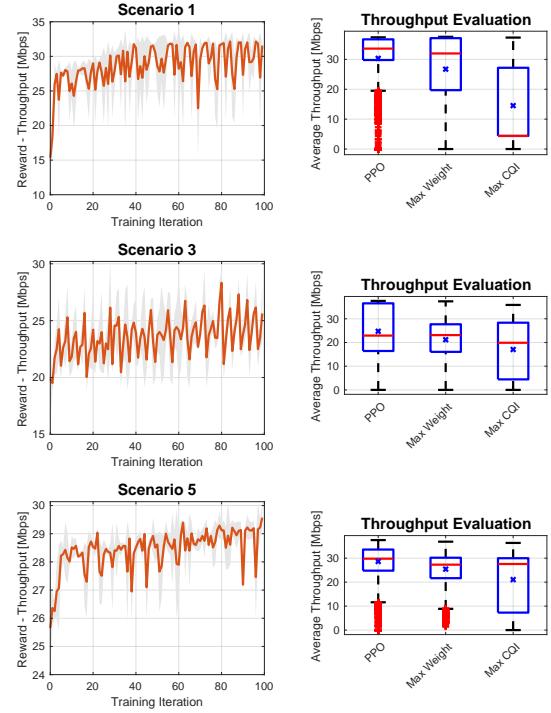


Figure 15: Can RL train and evaluate on DigitalTwin?

graphs in the first column illustrate the training curve for each scenario. We trained the policy network of RL for Scenario 1, 3 and 5 which are 2 Drone UEs, 2 Car UEs and 2 Drone UEs, and 2 Random Walk UEs and 2 Turntable UEs cases. We used specific CQI trace data to emulate channel conditions for individual scenarios. To train an RL agent using PPO, we designed a reward function to maximize throughput for

general scenarios. Each iteration, we sampled 5000 data samples including previous state, previous action and current state at each iteration and updated policy network until the reward saturates. Most transient periods for training were less than 40 iterations and the training curves converged for all scenarios. The graphs in the second column describe the throughput evaluation of the trained policy model. In each

graph, we compared the throughput of PPO to the ones of Max Weight and max CQI. In each algorithm, a red bar means the mean of its throughputs and a blue box their range. PPO obviously outperformed Max CQI, and even achieved almost the same throughput performance as Max Weight, which is usually considered as an optimum policy, with somehow higher ranges than Max Weight.