

Kairos: Energy-Efficient Radio Unit Control for O-RAN via Advanced Sleep Modes

J. Xavier Salvat Lozano*, Jose A. Ayala-Romero*, Andres Garcia-Saavedra* and Xavier Costa-Perez^{†‡}
NEC Laboratories Europe*, i2CAT Foundation[†], ICREA[‡]

Abstract—The high energy footprint of 5G base stations, particularly the radio units (RUs), poses a significant environmental and economic challenge. We introduce Kairos, a novel approach to maximize the energy-saving potential of O-RAN’s Advanced Sleep Modes (ASMs). Unlike state-of-the-art solutions, which often rely on complex ASM selection algorithms unsuitable for time-constrained base stations and fail to guarantee stringent QoS demands, Kairos offers a simple yet effective joint ASM selection and radio scheduling policy capable of real-time operation. This policy is then optimized using a data-driven algorithm within an xApp, which enables several key innovations: (i) a dimensionality-invariant encoder to handle variable input sizes (e.g., time-varying network slices), (ii) distributional critics to accurately model QoS metrics and ensure constraint satisfaction, and (iii) a single-actor-multiple-critic architecture to effectively manage multiple constraints. Through experimental analysis on a commercial RU and trace-driven simulations, we demonstrate Kairos’s potential to achieve energy reductions ranging between 15% and 72% while meeting QoS requirements, offering a practical solution for cost- and energy-efficient 5G networks.

Index Terms—5G, xApp, Advanced Sleep Modes, Radio Unit

I. INTRODUCTION

5G base stations (BS) consume four times more energy than their 4G counterparts [1]. This poses a serious environmental concern: with over 3.5 million 5G BSs deployed worldwide (and growing), the resulting carbon emissions caused by 5G BSs alone are estimated to exceed 100 *megatons* per year [2]. This also presents a financial challenge for mobile operators, potentially adding EUR 26.5 billion per year¹ to electricity bills, exacerbating already high operating expenses. Additionally, the higher power demand may necessitate retrofitting 30% of power supply systems, incurring an additional cost of approximately USD 2,800 *per site* in capital expenditure [4].

We make two key observations. First, this four-fold increase in power consumption is primarily attributed to the increased bandwidth, transmission power, and number of antennas that can be employed in 5G macro-cells. The radio unit (RU) accounts for 90% of this energy bill [1], which highlights a critical need to curb the RU’s energy drain in 5G macro-cells.

Second, while very few BSs experience zero traffic over longer time scales (tens of minutes), at shorter time resolutions (milliseconds), most cells are idle most of the time [5]. We confirm this in §II through our own measurements on cells from various operators in Madrid, Spain, in 2024. For illustration, Fig. 1 shows activity patterns for a few of those cells. In general, as we show in §II, the cumulative idle time amounts to over 50% of the total time in the median cell.

These two observations reveal substantial energy-saving potential through discontinuous transmissions (DTX), a fast

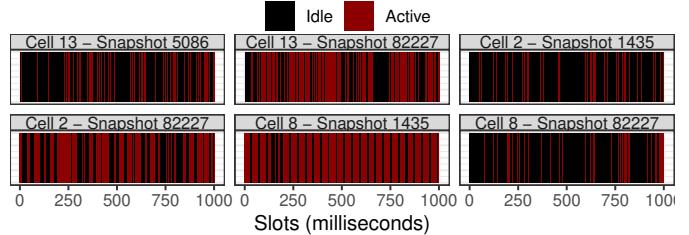


Fig. 1. Short and bursty energy-saving opportunities in production cells. Red/black slots are 1-ms periods with/without transmissions.

switching mechanism that enables deactivating the RU’s power amplifier—the most energy-consuming component, as we also show in §II—during these short, bursty idle periods [6]. Even a modest 1% reduction in RU consumption could potentially save over 1,200 GWh annually.

More recently, the O-RAN Alliance introduced specifications for controlling Advanced Sleep Modes (ASMs) in the RU, offering deeper sleep levels and greater energy savings through progressive deactivation of RU components [7]. However, deeper sleep levels come at the cost of increased switching overhead. Several research works have since proposed strategies to optimize the use of ASMs, such as [8], [9]. However, these works often make limiting assumptions about traffic models, neglect the computational complexity and real-time requirements of their solutions [10], and fail to provide hard QoS guarantees. A comprehensive review of related literature is presented in §VI.

To address these issues, we propose a different approach, which we call Kairos. Instead of relying on complex ASM control algorithms within the BS, which has real-time constraints, we advocate for a simple *joint ASM selection/radio scheduling policy* that is dynamically optimized in *near-real-time* based on context. This optimization is performed within an xApp in the O-RAN Near-Real-Time RAN Intelligent Controller (Near-RT RIC), with more relaxed timing requirements [11]. This allows us to design a data-driven algorithm with key innovations: (i) support for a *varying* number of network slices over time, with diverse/changing QoS requirements, and (ii) inciting joint ASM/radio policies that can minimize energy usage while providing *hard* QoS guarantees.

Kairos’s xApp-based approach has two key advantages: (i) it is suitable for deployment in systems with real-world constraints, as the computationally intensive optimization is offloaded to the Near-RT RIC, which has more relaxed timing constraints than the BS; and (ii) because the intelligence resides in an xApp and is observation-driven, Kairos can operate with various ASM selection strategies, including those proposed in [8], [9] or others reviewed in §VI. However, our results show that, when using Kairos, complex real-time ASM selection strategies do not yield substantial gains over simpler ones, such as the one we introduced in §III-A. This

¹Assuming 4.3 kW per 5G BS [1] and EUR 0.2008 per kWh of electricity costs for non-households in Europe [3].

is because, as our measurements on a commercial RU in §II demonstrate, consolidating radio resources into as few time slots as possible is sufficient for maximizing energy efficiency. Moreover, to provide hard QoS guarantees, the ASM strategy must be directly guided by the configured radio scheduling policy, an aspect that is neglected in the literature.

II. BACKGROUND & ANALYSIS

First, we provide background on 5G and O-RAN (§II-A), empirically study energy-saving opportunities in real-world networks (§II-B), analyze the power consumption profile of massive MIMO 5G radio units (§II-C), and detail relevant aspects of Advanced Sleep Modes and O-RAN control (§II-D).

A. A primer on 5G New Radio (NR) and O-RAN

NR defines the PHY/MAC procedures in 5G. We focus on sub-6GHz bands, which offer up to 100 MHz per band and flexible numerology $\mu = \{0, 1, 2\}$. The fundamental spectrum unit is the resource block (RB), consisting of 12 subcarriers spaced at $15 \cdot 2^\mu$ KHz. Time is divided into 1-ms subframes, each containing 2^μ slots carrying, typically, 14 OFDM symbols of duration $66.7 \cdot 2^{-\mu}$ μ s. In each Transmission Time Interval (TTI), often one slot, a scheduler assigns one transport block (TB) for each active User Equipment (UE). The TB size depends on the numerology, the buffered data, a RB/symbol scheduling policy, and the modulation and coding scheme (MCS), which depends on the signal-to-noise ratio (SNR).

To break vendor lock-in and enhance flexibility, the O-RAN Alliance has proposed a novel, open, and interoperable architecture for cellular networks with standardized interfaces [11]. Specifically, O-RAN establishes a 7-2x functional split between a Distributed Unit (DU) [11], which performs higher PHY functions (e.g., forward error correction), a Radio Unit (RU), in charge of lower PHY functions (e.g., FFT).

The O-RU and the O-DU are connected via the Fronthaul (FH) interface. FH traffic is logically divided into four planes: (i) the Control Plane (C-Plane), which transports real-time slot configuration control messages, (ii) the User Plane (U-Plane), which transports IQ samples data, (iii) the Synchronization Plane (S-Plane), which transports timing information messages and (iv) the Management Plane (M-Plane), which carries non-real-time management operation messages. C/U-Plane messages are exchanged using the eCPRI protocol which identifies them using the *eCPRI message type* field in the eCPRI header. eCPRI messages' payloads are structured using *sections*, which are data segments that carry information for each message type.

Finally, the Near-Real-Time RAN Intelligent Controller (Near-RT RIC) interacts with an O-RAN-compatible DU (O-DU) via the E2 interface, enabling the exchange of fine-grained performance measurements and radio scheduling policies at a ~ 100 ms granularity.

B. Small-timescale energy-saving opportunities

Since mobile systems are typically provisioned to handle peak traffic demands, individual cells often operate at low average utilization levels [12]. This inherent underutilization has motivated extensive research aimed at leveraging long-term

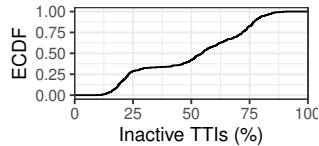


Fig. 2. Share of Idle TTIs in 10-minute periods across cells.

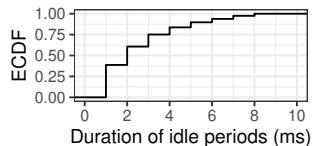


Fig. 3. Duration of idle periods across different cells.

traffic variations to improve energy efficiency through cell load consolidation and cell sleeping strategies. A comprehensive review of related work is provided in §VI.

However, workload patterns within individual cells reveal substantial power-saving sleeping opportunities at shorter timescales. To illustrate this, using Falcon [13], we collected millisecond-resolution workload data from diverse cells operated by multiple providers in Madrid, Spain, in May 2024. Consistent with prior research [12], the overall load is low (median less than 2 Mb/s) but highly variable (99th percentile around 10 Mb/s). For perspective, a 5G 100-MHz massive MIMO macro-cell can easily reach 1 Gb/s of peak capacity.

Moreover, we classified each TTI (1 ms in this case) as either "Active" (carrying traffic) or "Idle" (no traffic). Fig. 1 shows some activity patterns, *revealing the bursty nature of today's cellular traffic and highlighting potential energy savings through short-term sleep mechanisms*. Indeed, these short-term sleeping opportunities can amount to significant energy savings. Fig. 2 shows the distribution of the ratio of idle TTIs across all cells, calculated over 10-minute periods: the median inactivity is over 50%.

However, the duration of such idle periods is highly diverse. To illustrate this, Fig. 3 depicts their distribution across all monitored cells. The median idle period duration is 2 TTIs (28 OFDM symbols), but the 99th percentile reaches 8 TTIs (112 OFDM symbols). *This variability motivates the use of diverse sleep modes that can be implemented at intervals as short as one OFDM symbol or at longer intervals, with energy savings adjusted based on the specific case.* O-RAN refers to these modes as Advanced Sleep Modes (ASM)—see §II-D.

C. Radio Unit power consumption

We conducted an experimental analysis on a commercial 32TRX massive MIMO RU compliant with O-RAN 7.2-x FH specifications.² Based on these measurements, we can model the energy consumption of an RU as

$$P = P_0 + P_{RF} + P_{BB} + P_{PAM} \quad (1)$$

where

- P_0 models the power consumption of the baseline components in the RU (fabric, Ethernet, PCIe, etc.).
- P_{RF} models the power consumption of the analog-to-digital/digital-to-analog converters (ADC/DAC) and the frequency up/downconverters.
- P_{BB} models the power consumption of the core digital processor (e.g., an eFPGA) and the other low-PHY processing components (e.g., beamforming) for each antenna chain.
- P_{PAM} models the power consumption of the power amplifier module (PAM) at each antenna chain.

²Details omitted due to Non-Disclosure Agreement with the vendor.

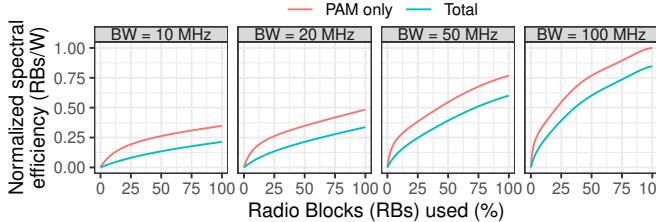


Fig. 4. Empirical results of spectral efficiency (RBs transmitted per watt of power consumed) of an O-RAN 7.2-x compliant 32TRX massive MIMO RU.

TABLE I
ADVANCED SLEEP MODES (ASM) CONSIDERED IN THIS PAPER.

| RU mode | Off RU HW (see §II-C) | Normalized power consumption | Switching delay (μ s) |
|---------|--------------------------|---------------------------------|-------------------------------|
| Idle | - | 1 | 0 |
| ASM 1 | PAM | 0.675 | 37 |
| ASM 2 | PAM, BB | 0.55 | 500 |
| ASM 3 | PAM, BB, RF | 0.23 | 5000 |

Our experiments suggest that P_0 , P_{RF} , P_{BB} remain rather static, independent of the load. However, P_{PAM} is highly dependent on the load and its consumption follows:

$$P_{PAM} = P_{PAM,0} \cdot P_{PAM,c}(r) \cdot \eta_{PAM} \cdot n_{ant} \quad (2)$$

where $P_{PAM,0}$ is a baseline power consumption that depends on the number of antennas, ganging elements, and the maximum EIRP per antenna; $P_{PAM,c}(r)$ is a factor that depends on the amount of subcarriers to be amplified, which in turn depends on the number of RBs r carried therein; η_{PAM} captures the efficiency of the amplifier to convert DC power into signal power gain and n_{ant} is the number of radio chains (with one amplifier per radio chain).

Fig. 4 presents experimental measurements on the RU configured with different bandwidths. The figure depicts the normalized spectral efficiency, measured as the number of radio blocks (RBs) transmitted per watt of power consumed in one symbol period. We present results for the complete RU (P in eq. 1) and for the PAM alone (P_{PAM} in eq. 1). As mentioned previously, the consumption of the remaining components is independent of the load. As expected, overall RU consumption is dominated by the PAM. Moreover, while the PAM's spectral efficiency plateaus as the utilized spectrum resources (RBs) increase, it is maximized when more bandwidth is employed. This observation motivates the principle behind the solution we propose in this paper: *consolidate data into as few symbol periods as possible, using more radio resources per period instead.*

D. 5G energy-saving features

3GPP has introduced a series of energy-saving features. Early on, 3GPP introduced a *lean carrier design* [14], which enables grouping together the BS's broadcast signals into periodic Synchronization Signal Blocks (SSB). The SSB periodicity can be adjusted by the operator in the range of 5 to 160 ms, to enable longer intervals for BS sleeping. Moreover, inherited from 4G, Discontinuous Transmission (DTX) [15] allows the transmitter to temporarily *sleep* when not sending data, conserving power during the aforementioned small timescale idle periods.

Advanced Sleep Modes (ASMs) enable progressively deeper sleep levels by deactivating more RU components, with the

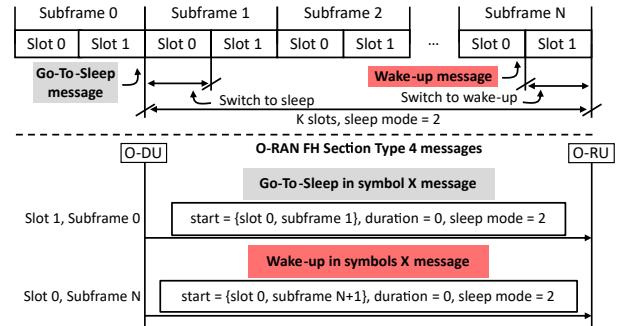


Fig. 5. Using O-RAN Section Type 4 messages to implement “Go-to-sleep” and “wake-up” commands. Example for ASM 2 and numerology $\mu = 1$.

trade-off of increased wake-up/sleep switching delays. Based on our analysis in §II-C and the related literature [16], we establish the ASMs described in Table I. Our RU already implements ASM 1 by deactivating the PAM, saving over 30% energy consumption compared to the active (yet idle) state. It requires one OFDM symbol period ($\sim 37 \mu$ s with numerology $\mu = 1$) for switching, aligning with existing DTX literature reporting PAM deactivation times of $30 - 65 \mu$ s [17], [18].

We simulate the remaining modes (ASM 2 and ASM 3) as follows. As shown in Table I, we assume that ASM 2 further deactivates the digital processor cores and other PHY layer components (BB in eq. 1). Consequently, we adopt a 0.5 ms switching delay, consistent with the literature [8], [16], [19], [20] and reported to C-state transitions in Intel processors [21]. Moreover, we assume that ASM 3 further deactivates RU components such as ADC/DAC and frequency converters (RF in eq. 1), with a 5 ms switching delay based on the literature [8], [16], [19], [20].

To asynchronously wake up or put the RU to sleep, Kairros relies on O-RAN’s FH C-Plane specification [22]. O-RAN enables the configuration of specific ASMs in an RU using “Section Type 4” C-Plane messages. These C-Plane messages convey slot-level configurations, including ASM settings, and apply to all endpoints associated with a carrier, antenna array, or the entire RU. Fig. 5 shows an example for numerology $\mu = 1$. To activate ASM 2, the DU sends the RU a “Go-to-sleep” command with a “Section Type 4” message with the following parameters: (i) a symbol mask indicating the symbol period within the next slot when the RU should enter sleep mode, (ii) “start” in the next immediate slot, (iii) duration set to 0 (undefined), and (iv) the ASM identifier. Conversely, to wake up the RU, the DU sends a “wake-up” message by unmasking the previously sent “Go-to-sleep” message.

III. KAIROS DESIGN

The analysis in §II highlights the diverse range of RU sleeping opportunities and the inherent trade-off between cell idle times and energy savings.

An additional observation is that 5G network slices have varying user delay requirements. For instance, while Ultra-Reliable Low-Latency Communication (URLLC) slices demand the lowest latency for critical applications, targeting delays of less than 1 ms, enhanced Mobile Broadband (eMBB) slices tolerate moderate delays for smooth streaming, aiming for 1-10 ms, or Massive Machine-Type Communica-

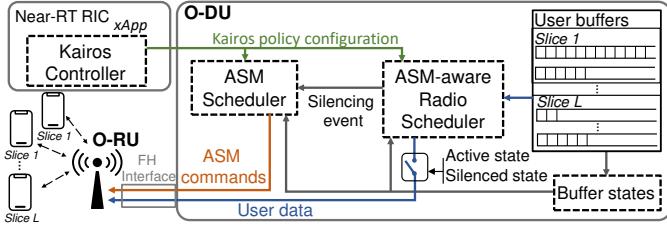


Fig. 6. Kairos system architecture.

tion (mMTC) slices, designed for massive IoT deployments, prioritize energy efficiency over strict latency requirements, often accepting delays of 10 ms and above. This observation presents an additional opportunity for energy savings: to trade off data delay for deeper and longer sleep intervals, *provided that slice QoS requirements are met*.

However, attaining such guarantees requires a radio scheduler at the MAC layer that is both RU-aware (to exploit sleeping opportunities) and QoS-driven. As we discuss in §VI, a large body of research studies schedulers with QoS guarantees (e.g., [23]), but ignore small-timescale RU energy-saving opportunities. Some recent work optimizes ASMs for this purpose, but they do not provide QoS guarantees (e.g., [16]). More importantly, very few of these works are actually practical for real-world base stations due to the tight computing constraints in the DU [10].

To address these challenges, we propose a novel approach called Kairos, which, unlike prior work, leverages the O-RAN-enabled interplay between real-time MAC-layer procedures (§III-A) and near-real-time radio-policing xApps (§III-B). This two-tier O-RAN-compliant framework is easy to implement and effectively exploits ASMs to save energy while guaranteeing hard QoS requirements. As depicted in Fig. 6, Kairos comprises three components: (i) the ASM-aware radio scheduler and (ii) the ASM scheduler, operating in real-time within the DU, and (iii) the Kairos controller, operating in near-real time within the Near-RT RIC, which are presented next.

A. Real-time DU operation: Scheduling and ASM selection

1) *ASM-Aware Radio Scheduler*: Our goal is to design a minimally intrusive radio scheduling policy that preserves the operator's intended scheduling logic and is lightweight enough to avoid violating MAC layer computing constraints, while aiding in extending RU sleep cycles within predefined QoS requirements. To this end, the Kairos policy establishes two states that a legacy MAC-layer scheduler must abide by:

- **Active**: The radio scheduler follows its own logic to allocate radio resources to users. However, the aggregated user data *must* be consolidated into as few OFDM symbols (time resources) as possible, maximizing the amount of RBs allocated per symbol period (spectrum resources). When the data buffer is empty, a silencing event is triggered.
- **Silenced**: The radio scheduler is not allowed to allocate radio resources, and downlink data is buffered. It also tracks the waiting time or “age” of the oldest user data burst in each slice, computed by the buffer states module (Fig. 6). If any burst’s age exceeds the Kairos policy configuration d , then an *activating* event is triggered. More formally, let \mathcal{L}

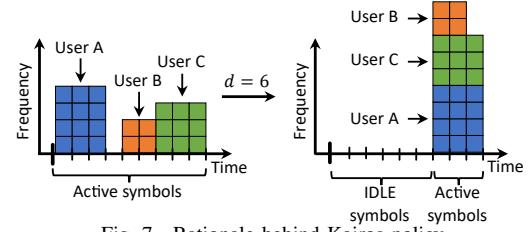


Fig. 7. Rationale behind Kairos policy.

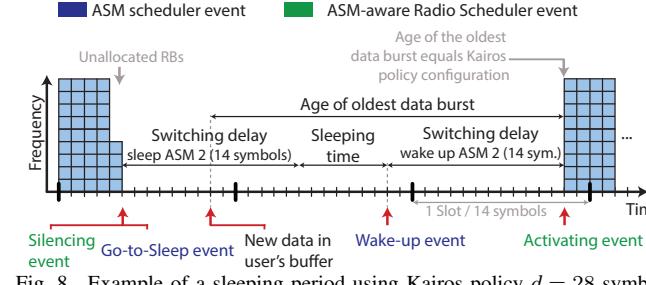


Fig. 8. Example of a sleeping period using Kairos policy $d = 28$ symbols.

be the set of network slices. We define $\mathcal{B}_t^{(l)}$ as the set of user data bursts from slice $l \in \mathcal{L}$ waiting in the downlink RLC buffers at time t . Let d_t be the Kairos policy configuration at time t . Thus, an *activating* event is triggered whenever $age(b) > d_t, \forall b \in \mathcal{B}_t^{(l)}, \forall l \in \mathcal{L}$.

Note that downlink transmissions shall resume in the symbol period immediately following an *activating* event, at which time all BS components should be awake. As shown in Fig. 7 for a toy example with a policy equal to $d = 6$ OFDM symbols, delaying small traffic bursts helps consolidating radio blocks into fewer OFDM symbols, which improves energy efficiency (see Fig. 4).

2) *ASM Scheduler*: The *ASM scheduler* generates ASM C-Plane commands based on some logic. Prior work has proposed somewhat complex ASM control algorithms (e.g., [16]). However, these approaches often overlook practical issues concerning computing and real-time constraints. Our goal is to devise a simple solution, suitable for real-time BS operation, that leverages the aforementioned policy to fully exploit the potential of ASMs for downlink transmissions.

When the ASM-aware radio scheduler triggers a silencing event, it notifies the ASM scheduler (Fig. 6). This component then selects an ASM based solely on the Kairos policy configuration. We propose a simple method: to select the deepest ASM such that the total latency of sleeping and waking up (see Table I) is lower than the Kairos policy configuration. In other words, the minimum duration of the sleep mode should fit within the silenced interval provided by the Kairos policy. For instance, in a scenario with a Kairos policy configuration of $d = 3$ ms, the selected sleep mode would be ASM 2. To configure the selected ASM in the RU, we use the O-RAN FH interface, as explained in §II-D. Note that the RU must be awake during the symbol period immediately following an activation event triggered by the radio scheduler. Therefore, the ASM scheduler must notify the RU to wake up in advance, with the lead time depending on the specific sleep mode.

Fig. 8 illustrates the operation of the Kairos policy, where the x-axis represents time resources (symbol periods and slots), and the y-axis depicts radio resources (RBs). Unused

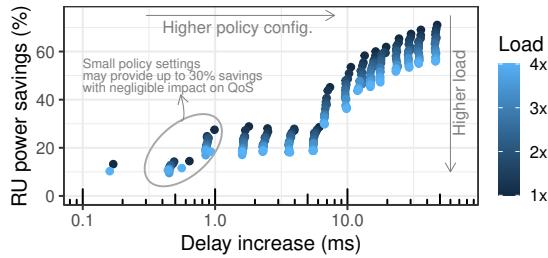


Fig. 9. Analysis of Kairos radio scheduling policy.

RBs in the fifth symbol trigger a “Silencing” event in the radio scheduler, leading to a “Go-to-sleep” command in the ASM scheduler. Since the Kairos policy is configured with $d = 1$ ms (2 slots), ASM 2 is selected because it is the deepest mode with a total switching delay less than d (see Table I). Seven symbol periods later, new data arrives while the RU is asleep. After 2 slots, the age of this data burst shall meet the policy configuration, which should cause the radio scheduler to trigger an “Activating” event. Therefore, a “wake-up” message is scheduled one slot before the “Activating” event, aligning with ASM 2’s one-slot switching delay.

3) *Analysis:* In this way, the Kairos policy enables selecting a Pareto-optimal operating point tailored to the requirements of a given network slice (policy optimization is addressed later). Fig. 9 illustrates the trade-off between RU power savings (y-axis) and extra delay incurred (x-axis) for various Kairos policy configurations and load levels. We adopt a trace-driven approach, using the data from real cells that we presented in §II, to emulate load on the same O-RAN 7.2-x compliant 32TRX massive MIMO RU employed in §II-C, configured to use 50 MHz of bandwidth and numerology $\mu = 1$ (see §II-A). We define “1x” load as the load patterns observed in our measurements. To ensure our analysis remains relevant for potential future scenarios with higher cell loads, we generate these higher loads by virtually compressing time in our traces by a factor indicated in the color legend (“2x”, “3x”, and so on). This method allows us to analyze diverse load levels while preserving the bursty nature of real-world traffic patterns.

The results offer two key insights. First, small policy settings that induce minimal delay increases (0.25-1 ms) enable RU power savings of 30% (under today’s typical loads, i.e., “1x”) to 10% (under loads four times higher than the median cell load today, i.e., “4x”). Second, if QoS requirements were more relaxed, permitting 10-40 ms of additional user delay, the Kairos policy could achieve even greater savings, ranging from 70% (under “1x” loads) to 35% (under higher loads).

B. Near-RT RIC xApp: Kairos controller

The *Kairos Controller* is essentially an O-RAN xApp running at the Near-RT RIC that dynamically computes the policy configuration d (buffering delay threshold) over time. As shown in [24], different services (hence, slices) show different temporal load patterns even if they belong to the same category. Hence, it is of paramount importance to dynamically adjust the policy configuration, considering the behavior of all the slices, in order to harness the maximum possible energy savings while guaranteeing QoS requirements. This is a challenging task as the user’s experienced delay does not only

depend on the Kairos policy but also on the traffic profile that the BS has to handle (i.e., traffic load and type of other users, other network slices, etc). We address this in detail in §IV.

IV. KAIROS CONTROLLER

Next, we first formally describe the policy control problem (§IV-A) and then propose a solution (§IV-B).

A. Problem formulation

As explained above, the Kairos controller is responsible for adjusting the Kairos policy configuration used by the real-time modules (ASM-aware radio scheduler and ASM scheduler) based on changes in the traffic demands and different QoS requirements across slices. Following the O-RAN architecture, the Kairos controller operates in the Near-RT RIC, i.e., with a time granularity of ~ 100 ms. Hence, we define the time steps denoted by $t = \{0, 1, \dots\}$ based on this time granularity.

We let $d_t \in \mathcal{D}$ denote the Kairos policy configuration at time step t . We do not allow different policies for different slices for two reasons: (i) learning a separate policy per slice from data would exponentially increase the time and data required for convergence to a good model; and (ii) although different slices may have different user delay requirements, the physical RU is shared. Thus, if the policy for one slice triggers the scheduler into an “Active” state, there is no energy-saving benefit from deferring transmissions on other slices, as the RU would already be active.

Furthermore, we consider a BS with a set \mathcal{L} of L_{\max} network slices. We do not make any assumption on the distribution of users across slices. Given that the number of active slices in the network can change over time, we denote by $\mathcal{L}_t \subseteq \mathcal{L}$ the set of active slices at time t , whose cardinality is $|\mathcal{L}_t| = L_t$. The target requirement of slice $l \in \mathcal{L}$ is denoted by $\Delta^{(l)}$.

We now let $\mathcal{I}_t^{(l)}$ describe the set of data bursts generated by the users in slice $l \in \mathcal{L}$ during the time step t . Each data burst $I \in \mathcal{I}_t^{(l)}$ is defined by its size (in bits) and the generation time. We characterize the user data bursts using the distribution of their inter-arrival time (IAT) and size. Specifically, for a given set $\mathcal{T} \in [0, 1]^N$ of N quantiles, we define $\zeta_t^{(l)} = \Phi^{\text{IAT}}(\mathcal{I}_t^{(l)}, \mathcal{T})$ as the vector of quantile values of the IAT distribution. Similarly, $\xi_t^{(l)} = \Phi^{\text{size}}(\mathcal{I}_t^{(l)}, \mathcal{T})$ denotes the vector of quantile values of the burst size distribution. Then, we define the context of slice l at time t as $s_t^{(l)} := (\zeta_t^{(l)}, \xi_t^{(l)})$. Now, we can define the joint context of the BS by $s_t = \{s_t^{(l)}\}_{l \in \mathcal{L}_t} \in \mathcal{S}$.

We now define the energy consumption of the system as $E(s_t, d_t)$. Since all slices operate over the same RU, the consumed energy is a global metric. We define the QoS metric per slice as $\delta^{(l)}(s_t, d_t)$. For example, this metric can be related to the delay experienced by users. Finally, we define the function $\pi : \mathcal{S} \mapsto \mathcal{D}$ that maps joint contexts to Kairos policy configurations. Consequently, our objective is to solve the following contextual bandit problem:

Problem 1:

$$\begin{aligned} & \underset{\pi}{\operatorname{argmin}} \sum_{t=1}^T E(s_t, \pi(s_t)) \\ \text{s.t. } & \delta^{(l)}(s_t, \pi(s_t)) < \Delta^{(l)}, \quad \forall l \in \mathcal{L}_t, t = 1, \dots, T. \end{aligned}$$

Problem 1 aims to minimize the average energy consumption of the RU while satisfying the *hard* QoS constraints of every slice at each time step.

We would like to emphasize that the problem in (1) is challenging due to several factors. Both the energy $E(\cdot)$ and the QoS per slice $\delta^{(l)}(\cdot)$ are complex functions that depend not only on s_t and $\pi(\cdot)$, but also on other system parameters (e.g., bandwidth). Intuitively, higher values of d_t can reduce power consumption by enabling longer silent periods. However, the value of the QoS metric $\delta^{(l)}(\cdot)$ is complex to devise in advance as it depends on the traffic and buffer states of all other slices. Moreover, these metrics can exhibit randomness due to unknown and uncontrollable environmental factors, which must be considered to satisfy the hard constraints. Furthermore, the number of active slices can vary over time, increasing the problem's complexity and precluding the use of standard machine learning approaches (with fixed input size and number of constraints). We next overcome these challenges with a novel learning framework.

B. A novel ML framework for ASM control

Our solution is designed to overcome the next challenges:

- *Variable input size*: The size of the context s_t may change over time due to the varying number of active slices. Hence, we propose a dimensionality-invariant encoder that projects the context into a fixed dimensional space.
- *Hard constraints with uncertain QoS*: Problem (1) imposes hard constraints that are difficult to satisfy, as QoS measurements in the network $\delta^{(l)}(\cdot)$ can exhibit randomness due to unknown and uncontrollable factors. To address this, we use distributional critics to capture not only the mean value of the function $\delta^{(l)}(\cdot)$ but also its distribution. Thus, by examining the tail of the distribution, we can ensure constraint satisfaction with some pre-determined probability.
- *Variable number of constraints*: The problem involves a potentially large and variable number of constraints. To address this, we propose an ML architecture with one actor (to select d_t) and multiple distributional critics (one per constraint). This model allows us to compute a joint cost function and backpropagate the loss through all active critics (based on the number of active slices), with the goal of learning the optimal actor function.

In the following, we detail each component of our solution.

1) *Dimensionality-invariant encoder*: As mentioned previously, the number of active slices L_t can be potentially large and vary over time. A naive approach to inputting s_t into the machine learning framework would be to concatenate $s_t^{(l)}$ $\forall l \in \mathcal{L}_t$ and then use a single neural network (NN) with the appropriate input size. However, this requires a separate NN for each possible combination of active slices, totalling $\sum_i {L_{\max}}^i$ encoders, as the input is not permutation invariant. In other words, the traffic of each slice l is associated with a specific target QoS $\Delta^{(l)}$.

Instead, we propose a much simpler, scalable, and effective solution inspired by relational networks [25] and graph neural networks [26]. We project the context of each slice into a

higher-dimensional space and then combine the projections of all slices using an aggregator function. While summation is a common choice for aggregation in the literature, it is permutation invariant. To differentiate between contexts corresponding to different slices, we incorporate the slice identifier into the projection function. Thus, the encoded context representation with a fixed number of dimensions is given by:

$$\tilde{s}_t = \sum_{l \in \mathcal{L}_t} g(s_t^{(l)}, l|\phi), \quad (3)$$

where $g(\cdot|\phi)$ is the projection function with parameters ϕ .

2) *Distributional critics*: The main idea behind our distributional critics is to capture the distribution (i.e., the value of a set of quantiles) of the target function. This enables us to precisely identify the tail of the distribution and thus satisfy constraints with certain probability.

For a given random variable X , the value of its quantile function is defined as $q_\tau = F_X^{-1}(\tau)$, where $\tau \in [0, 1]$ is the quantile and $F_X(x)$ is the cumulative distribution function. We employ the quantile regression loss, an asymmetric convex function that penalizes underestimation error with weight $1-\tau$ and overestimation error with weight τ :

$$\mathcal{J}^\tau(\hat{q}_\tau) := \mathbb{E}_{x \sim X} [\rho_\tau(x - \hat{q}_\tau)], \text{ where} \quad (4)$$

$$\rho_\tau(u) := u \cdot (\tau - \delta_{\{u < 0\}}) \quad \forall u \in \mathbb{R}, \quad (5)$$

where \hat{q}_τ is the estimated value of the quantile function, and $\delta_{\{z\}}$ is an indicator function that takes the value 1 when z holds and 0 otherwise.

For each critic, we approximate N quantile values $q_{\tau_1}, \dots, q_{\tau_N}$ by training the critic using gradient decent to minimize the following objective:

$$\sum_{i=1}^N \mathcal{J}^{\tau_i}(\hat{q}_{\tau_i}). \quad (6)$$

Since the loss function defined above is not smooth at $u = 0$, the performance of function approximators (e.g., NNs) may not be optimal. To address this, we use the *quantile Huber loss* [27]. This loss function, instead, presents a squared shape in the interval $[-\kappa, \kappa]$, and reverts to the standard quantile loss outside this interval:

$$J_\kappa(u) := \begin{cases} \frac{1}{2}u^2 & \text{if } |u| \leq \kappa \\ \kappa(|u| - \frac{1}{2}\kappa) & \text{otherwise.} \end{cases} \quad (7)$$

The asymmetric version of the Huber loss is given by:

$$\rho_\tau^\kappa(u) := |\tau - \delta_{\{u < 0\}}| \frac{J_\kappa(u)}{\kappa}. \quad (8)$$

By substituting $\rho_\tau^\kappa(u)$ into eq (4), we obtain the quantile Huber loss. Note that as $\kappa \rightarrow 0$, the quantile Huber loss reverts to the quantile regression loss.

3) *Framework architecture and learning procedure*: To handle a large and variable number of constraints in our problem, we use an architecture with one actor and $L_{\max} + 1$ distributional critics [28].

The actor function $\pi(\tilde{s}|\eta)$, parameterized by η , can take continuous values as the Kairos policy is continuous. This type

of actor is commonly referred to as a deterministic actor [29]. The distributional critics are denoted by $C^{(l)}(\tilde{s}, d | \theta^{(l)})$, where $\theta^{(l)}$ represents the parameters of critic l . We use the index $l = 0$ for the critic that approximates the objective function (energy consumption in (1)) and $l = 1, \dots, L_{\max}$ for the critics that approximate the constraint of slice l , i.e., $\delta^{(l)}(\cdot)$.

We define an aggregate cost signal to capture the information about all constraints provided by the critics:

$$C_t^{agg}(\tilde{s}_t, d_t, | \theta) := \bar{C}^{(0)}(\tilde{s}_t, d_t | \theta^{(0)}) + \sum_{l \in \mathcal{L}_t} \lambda \max \left(\gamma^\alpha(C^{(l)}(\tilde{s}_t, d_t | \theta^{(l)})) - \Delta^{(l)}, 0 \right) \quad (9)$$

where $\gamma^\alpha(X)$ is the quantile function value of distribution X at quantile α , λ is the penalty weight for the constraints, $\theta = \{\theta^{(0)}, \dots, \theta^{(L_{\max})}\}$ is the joint set of parameters of the $L_{\max} + 1$ critics, and $\bar{C}^{(l)}(\cdot)$ is the mean of the distribution of critic l .

In detail, the first term in eq. (9) represents the mean of the objective function (energy) that we aim to minimize, while the second term aggregates the penalty incurred by the constraint violations. When a constraint is satisfied, the value inside the $\max()$ function is negative and the penalty is zero. Importantly, the use of the quantile function in the aggregated cost function ensures that the tail of the distribution of the constraints (as $\alpha \rightarrow 1$) meets the constraints, boosting the robustness of our solution in terms of constraint satisfaction. While this formulation focuses on QoS metrics that should be below a maximum value (e.g., delay), eq. (9) can be easily adapted for QoS metrics that have a minimum target (e.g., throughput) by selecting values of α close to zero and reversing the sign inside the maximum operator.

To train the actor, we first define its objective function as

$$R(\pi) = \mathbb{E}_{\tilde{s} \sim \beta}[C^{agg}(\tilde{s}, \pi(\tilde{s} | \eta) | \theta)] \quad (10)$$

where $\beta(\tilde{s})$ is the stationary distribution of the projected contexts. Note that in a contextual bandit problem, the distribution of the contexts is not conditioned by the actor function.

Then, we derive the actor update by applying the chain rule to the performance objective defined in eq. (10) with respect to the actor parameters [29]:

$$\nabla_\eta R(\pi) \approx \mathbb{E}_{\tilde{s} \sim \beta} [\nabla_a C^{agg}(\tilde{s}, d | \theta) |_{d=\pi(\tilde{s} | \eta)} \nabla_\eta \pi(\tilde{s} | \eta)]. \quad (11)$$

For a practical implementation of the algorithm, we also consider a replay buffer \mathcal{D} to store samples of experience from each time step, as in [30]. Gradients are then computed using mini-batches of B samples randomly drawn from the replay buffer. Empirically, we observed that the distributional critics achieve better performance when approximating multiple quantiles rather than just α . Therefore, we define T as the set of quantiles approximated by the critics, where $\alpha \in T$. At each time step t , we add noise \mathcal{N}_t to the output of the actor to promote exploration during training [31]. Fig. 10 illustrates all the components of the Kairos controller and Algorithm 1 presents its training procedure.

In this way, Kairos is purposely designed to maximize energy efficiency, consolidating radio resources in as few OFDM

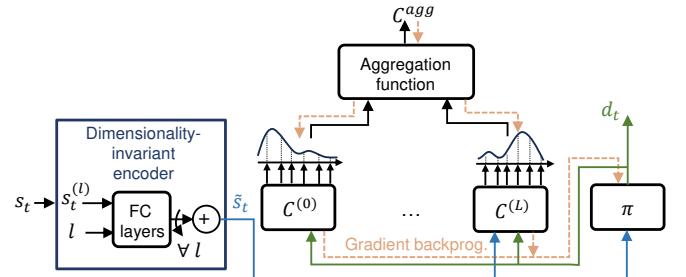


Fig. 10. Kairos controller comprising a dimensionality-invariant encoder, a deterministic actor, $L + 1$ critics, and an aggregation function (eq. (9)).

Algorithm 1 Kairos controller training

```

Input:  $\alpha, \kappa, B, T$ 
Initialize: Parameters  $\eta, \theta$ ; reply buffer  $\mathcal{D} = \emptyset$ ;
Exploration noise  $\mathcal{N}$ ;
1: for  $t = 1, \dots, T$  do
2:   Observe the context of all active slices  $s_t$ 
3:   Compute  $\tilde{s}_t$  using eq. (3)
4:   Compute the Kairos policy configuration  $d_t = \pi(\tilde{s}_t, | \eta) + \mathcal{N}_t$ 
5:   Apply the Kairos policy configuration during the time step  $t$ 
6:   At the end to step  $t$ , observe the energy  $E_t$ 
7:   Observe the QoS per slice  $\delta_t = \{\delta_t^{(l)}\}_{l \in \mathcal{L}_t}$ 
8:   Store in  $\mathcal{D}$  the experience sample  $\langle \tilde{s}_t, d_t, E_t, \delta_t \rangle$ 
9:   Sample a random minibatch of  $B$  samples  $\langle \tilde{s}_i, d_i, E_i, \delta_i \rangle$ 
10:  Update the critic 0 (energy) by minimizing the loss
     $\frac{1}{B} \sum_i \sum_{\tau \in T} \rho_\tau^\kappa(E_i - \gamma^\tau(C^{(0)}(\tilde{s}_i, d_i | \theta^{(0)})))$ 
11:  for  $l \in \mathcal{L}_t$  do
12:    Update the critic  $l$  by minimizing the loss
     $\frac{1}{B} \sum_i \sum_{\tau \in T} \rho_\tau^\kappa(\delta_i^{(l)} - \gamma^\tau(C^{(l)}(\tilde{s}_i, d_i | \theta^{(l)})))$ 
13:  end for
14:  Update the actor with the sampled policy gradient
     $\frac{1}{B} \sum_i \nabla_d C^{agg}(\tilde{s}_i, d, | \theta) |_{d=\pi(\tilde{s}_i | \eta)} \nabla_\eta \pi(\tilde{s}_i | \eta)$ 
15: end for

```

symbols as possible, within a deterministic delay bound. This is in contrast with the literature's ASM-only strategies, which are unable to provide hard QoS guarantees.

V. PERFORMANCE EVALUATION

To evaluate Kairos, we utilize the commercial RU and the emulator of real-world traces introduced in §II. Using this testbed, we implement Kairos as described in §III and §IV, and evaluate its performance in terms of convergence (§V-A), in comparison with alternative radio and ASM scheduling solutions (§V-B), and its performance in dynamic environments with time-varying network slices (§V-C).

Moreover, Kairos is configured with $\alpha = 0.995$, a reply buffer \mathcal{D} of 10^4 samples, and batch size $B = 128$. The training exploration noise \mathcal{N} follows an Ornstein-Uhlenbeck process with parameters $\theta_{\text{noise}} = 0.15$ and $\sigma_{\text{noise}} = 0.15$ [31]. Finally, we consider a decision time step of 200 ms.

A. Convergence

To study the convergence of Kairos, we test the traces from §II under various load levels, starting without prior training. We repeat the experiment for different maximum UE delay QoS requirements. Fig. 11 depicts the evolution over time of the mean user delay (top) and normalized RU power consumption³ (bottom) for each scenario.

³Power consumption figures are normalized to respect the Non-Disclosure Agreement with equipment vendor.

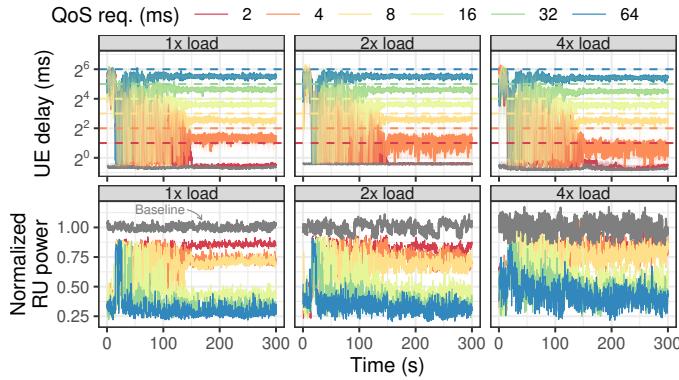


Fig. 11. Kairos convergence. The dashed lines represent the QoS requirement. The grey lines represent the performance of an ASM-unaware baseline.

Kairos consistently meets the QoS target after approximately 150 seconds of training (750 time steps). To this end, Kairos induces higher RU power consumption for stricter QoS requirements. For example, under a “4x” load, the mean power consumption is 0.39 when $\Delta = 64$ ms, but increases to 0.85 (over twofold) when $\Delta = 2$ ms. This is expected, as stricter QoS targets require less spectrally efficient allocations as those discussed in §II-C. However, our approach achieves power savings that range between 72% (“1x” load and $\Delta = 64$ ms) and 15% (“4x” load and $\Delta = 2$ ms) compared to an ASM-unaware baseline (depicted in grey).

We observe that user delay, our measure of QoS performance, more closely approaches its target as the target becomes less strict. This is because the bursty traffic behavior of our traces smoothens out with higher Kairos policy settings, as more data bursts are buffered. This behavior increases traffic predictability and allows Kairos to operate more aggressively, saving energy while still meeting the QoS target. Conversely, tighter QoS targets limit the radio scheduler’s ability to buffer and smooth out data bursts, requiring a more conservative operation to ensure QoS guarantees are met.

B. Comparison

As we show in §VI, there exists recent work on ASM control. However, they are unable to provide hard QoS guarantees, and most cannot operate in real time. Hence, we next compare Kairos with different benchmarks that can guarantee QoS.

1) *ASM scheduler*: We first evaluate Kairos’s ASM scheduler against an oracle, both operated by Kairos controller. This oracle ASM scheduler is an idealized approach with perfect foresight of future data burst arrivals, allowing it to make the optimal ASM schedule. Fig. 12 depicts the power consumption achieved by both approaches for a “4x” load scenario (other load levels yield similar results). Notably, the figure demonstrates that Kairos, a simple strategy suitable for real-time operation, performs comparably to the oracle when guided by a near-real-time policy. This contrasts sharply with related, more complex solutions [8], [16], [32], which are difficult to implement in real systems due to their computational complexity and the tight computing constraints of DUs.

2) *Kairos controller*: We compare Kairos controller against two alternative ML architectures that solve our problem:



Fig. 12. Kairos ASM scheduler vs an Oracle solution for “4x” load.

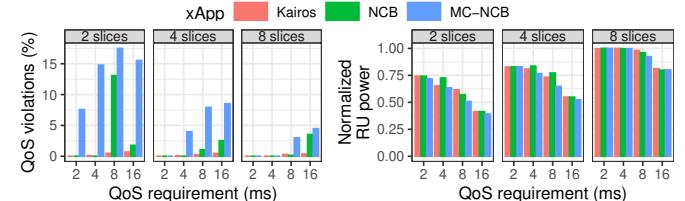


Fig. 13. Comparison between different xApp ML solutions. Load = “1x”.

- **NCB** (Neural Contextual Bandit) comprises an actor-critic NN architecture [31] modified to fit our problem formulation (see [28]). We define a single utility function where the constraints are Lagrangian-like penalty terms.
- **MC-NCB** (Multi Critic-NCB) extends NCB with one critic per constraint. Each (non-distributional) critic approximates the constraint function using MSE loss. The actor is updated, and the output of all critics is aggregated similarly to eq. (9).

Fig. 13 depicts the share of data bursts violating their delay target (left plot) and normalized RU power consumption (right plot) for different delay targets (x-axis) and number of slices (2, 4, and 8, respectively), each following the “1x” load pattern. As expected, RU power consumption decreases with relaxed QoS requirements and a lower number of slices (i.e., lower aggregated load). However, while RU power consumption is similar across all three approaches, Kairos significantly outperforms the other two benchmarks in terms of QoS violation rates, achieving nearly negligible violations. Notably, QoS violations are higher with fewer slices (lower aggregated load). This occurs because increasing the load smooths the traffic burstiness, making it more predictable and creating a friendlier environment for the controllers.

C. Dynamic environments

We conclude by evaluating Kairos and the two benchmarks in a more dynamic and heterogeneous setup, emulating a scenario with five network slices (each with distinct QoS requirements of 16, 8, 4, 2, and 1 ms) handling “1x” load patterns and joining the system every 30 seconds. All ML models are trained identically. Fig. 14 shows the burst delay performance (top), excess delay relative to QoS targets (middle), and normalized power consumption (bottom) over time.

All ML models achieve comparable power savings compared to the baseline, with diminishing returns as the number of slices and aggregate load increase, and QoS requirements become stricter (1 ms for slice 5). However, only Kairos consistently maintains negligible QoS violations across all system states. In contrast, NCB and MC-NCB exhibit significant violations, reaching 19.7% and 51% at the 99th percentile, respectively. These violations are particularly pronounced during

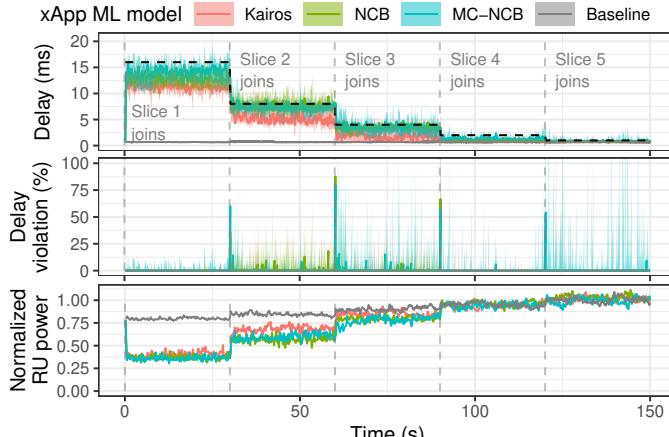


Fig. 14. 5 slices with QoS requirements $d = \{16, 8, 4, 2, 1\}$ ms, respectively, joining the system every 30 seconds. The dashed line represents the most stringent QoS requirement. The grey lines represent the performance of an ASM-unaware baseline.

state transitions as new slices join, whereas Kairos seamlessly adapts to new slices without incurring additional delays.

VI. RELATED WORK

The concept of cell sleeping for energy saving has been extensively explored in the literature [33]. For instance, the authors in [34] address the joint problem of cell clustering and BS sleeping, while [35] proposes a joint strategy for cell sleeping and interference coordination. These works exploit medium- to long-term traffic fluctuations to consolidate load into fewer active cells through coordination.

However, as demonstrated in this and other works, small-timescale cell sleeping opportunities exist to save energy in individual cells without requiring cell coordination. As explained in §II, Discontinuous Transmissions (DTX) [15] allow the transmitter to deactivate hardware components (e.g., power amplifier in the RU) during idle periods with minimal overhead, conserving power. DTX has also been explored in the literature; for example, [36] derives a hysteresis-based sleeping strategy considering the maximum allowable delay.

While ASMs [22] offer the potential for greater energy savings than DTX alone, they may also increase user delay due to transmissions potentially being deferred until the BS is fully operational. Several studies have proposed strategies to optimize ASM usage [8], [9], [16], [19], [20], [32], [37]. However, these previous works have significant limitations, hindering their integration into real-world systems:

Suboptimal ASM scheduling strategies. There is no one-size-fits-all solution that can accommodate the diverse range of real-world bursty traffic conditions. For instance, some approaches, such as those in [8], [16], propose transitioning from the deepest sleep mode (ASM 3) through intermediate sleep modes (ASM 2, ASM 1) before waking up during an idle cycle. In contrast, others like [20] suggest starting with ASM 1 and progressively entering deeper modes until ASM 3 is reached before waking up. Both strategies employ elaborate schemes (e.g., data-driven models) to optimize the duration of each sleep mode, aiming to balance energy savings against the extra delay caused by mode transitions. However,

these strategies can be suboptimal under realistic bursty traffic conditions: the former may overuse deep sleep modes in high or medium-load periods, while the latter may do so under low load conditions. As demonstrated in this paper, simpler ASM scheduling strategies provide similar performance.

Unrealistic assumptions. Many of these works rely on unrealistic assumptions, such as Poisson or log-normal traffic arrivals [8], [16], [32]. While these models are suitable for long-term traffic behavior, they fail to capture the bursty nature of traffic at millisecond-level granularity (see §II). Therefore, it is key to validate ASM strategies using realistic traffic models or, as in our paper, traces collected from real base stations.

Lack of hard QoS guarantees. Clearly, ASM utilization entails a trade-off between energy savings and delay. However, most related works do not consider hard QoS constraints, opting instead to approximate the constrained problem using Lagrange-like utility functions suitable for data-driven models [8], [9], [16], [32], [37]. While acknowledging this issue, [9] proposes a risk model to mitigate scenarios where traffic exists but the BS is sleeping. However, varying delay requirements across different applications, services, or network slices motivate sacrificing delay for further energy savings, provided these requirements are met. This necessitates a solution with hard delay guarantees, which, to our knowledge, our paper is the first to provide in this context.

Compliance with real system constraints. All the previous works propose performing ASM scheduling, which is a real-time (sub-ms) operation in the BS, based on rather complex methods (most of which based on reinforcement learning models) [8], [9], [16], [19], [20], [32], [37], which are hard to implement under such computing constraints [10]. As we showed in this paper, simpler ASM scheduling policies, suitable for real-time operation in the BS, are sufficient as long as joint ASM/radio policies optimized in near-real-timescales are in place, which is what Kairos achieves.

VII. CONCLUSIONS

In contrast to related work that proposes complex ASM optimization algorithms for energy savings in the Radio Unit (RU) of 5G base stations, this paper introduces Kairos, a novel approach leveraging simple, real-time ASM and radio scheduling policies jointly optimized in near-real time by an O-RAN xApp. Kairos's unique strength lies in a novel ML framework that dynamically adapts to varying network conditions, such as the addition or removal of network slices, while maintaining pre-defined QoS requirements. Through rigorous evaluation using real-world traffic patterns and a commercial RU, we demonstrate Kairos's ability to achieve substantial energy reductions between 15% and 72% in the RU while preserving QoS constraints, establishing it as a practical solution for cost- and energy-efficient 5G networks.

VIII. ACKNOWLEDGMENTS

Work supported by the European Commission through grants No. 101139270 (ORIGAMI) and SNS-JU-101097083 (BeGREEN), by NextGeneration EU through UNICO I+D grant no. TSI-063000-2021 (OPEN6G), and by the CERCA Programme.

REFERENCES

- [1] S. Han, S. Bian *et al.*, "Energy-efficient 5G for a greener future," *Nature Electronics*, vol. 3, no. 4, pp. 182–184, 2020.
- [2] T. Li, L. Yu, Y. Ma, T. Duan, W. Huang, Y. Zhou, D. Jin, Y. Li, and T. Jiang, "Carbon emissions and sustainability of launching 5G mobile networks in China," 2023. [Online]. Available: <https://arxiv.org/abs/2306.08337>
- [3] Eurostat. (2024) Electricity price statistics: Electricity prices for non-household consumers. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Electricity_price_statistics#Electricity_prices_for_non-household_consumers.
- [4] Huawei. (2020) 5G Power: Creating a green grid that slashes costs, emissions & energy use 5G Power: Creating a green grid that slashes costs, emissions & energy use. <https://www.huawei.com/en/huaweitech/publication/89/5g-power-green-grid-slashes-costs-emissions-energy-use>.
- [5] F. Parzysz and Y. Gourhant, "Drastic Energy Reduction With gDTX in Low Cost 5G Networks," *IEEE Access*, vol. 6, pp. 58 171–58 181, 2018.
- [6] P. Frenger, P. Moberg, J. Malmodin, Y. Jading, and I. Godor, "Reducing Energy Consumption in LTE with Cell DTX," in *2011 IEEE 73rd Vehicular Technology Conference (VTC Spring)*, 2011, pp. 1–5.
- [7] O-RAN Alliance, "Security Requirements and Controls for O-RAN Defined Interfaces and Network Functions," O-RAN Alliance, Technical Specification O-RAN.WG4.CUS.0-R003-v13.00, 2023.
- [8] F. E. Salem, T. Chahed, Z. Altman, and A. Gati, "Traffic-aware advanced sleep modes management in 5G networks," in *2019 IEEE wireless communications and networking conference (WCNC)*. IEEE, 2019, pp. 1–6.
- [9] M. Masoudi, E. Soroush, J. Zander, and C. Cavdar, "Digital twin assisted risk-aware sleep mode management using deep Q-networks," *IEEE Transactions on Vehicular Technology*, vol. 72, no. 1, pp. 1224–1239, 2022.
- [10] Y. Huang, S. Li, Y. T. Hou, and W. Lou, "GPF+: A novel ultrafast GPU-based proportional fair scheduler for 5G NR," *IEEE/ACM Transactions on Networking*, vol. 30, no. 2, pp. 601–615, 2021.
- [11] A. Garcia-Saavedra and X. Costa-Perez, "O-RAN: Disrupting the virtualized ran ecosystem," *IEEE Communications Standards Magazine*, 2021.
- [12] L. L. Schiavo *et al.*, "CloudRIC: Open Radio Access Network (O-RAN) Virtualization with Shared Heterogeneous Computing," in *Proceedings of the 30th Annual International Conference on Mobile Computing and Networking*, ser. ACM MobiCom '24. New York, NY, USA: Association for Computing Machinery, 2024, p. 558–572.
- [13] R. Falkenberg and C. Wietfeld, "FALCON: An accurate real-time monitor for client-based mobile network data analytics," in *2019 IEEE Global Communications Conference (GLOBECOM)*. Waikoloa, Hawaii, USA: IEEE, Dec. 2019. [Online]. Available: <https://arxiv.org/abs/1907.10110>
- [14] E. Dahlman, S. Parkvall, and J. Skold, *5G NR: The next generation wireless access technology*. Academic Press, 2020.
- [15] 3GPP, "Nr; medium access control (mac) protocol specification," 3GPP, Technical Specification 38.321, 2024, accessed: January 10, 2025. [Online]. Available: [invalidURLremoved]
- [16] A. A. Razzac, T. Chahed, Z. Shamseddine, and W. Zahwa, "Advanced Sleep Modes in 5G Multiple Base Stations Using Non-Cooperative Multi-Agent Reinforcement Learning," in *GLOBECOM 2023-2023 IEEE Global Communications Conference*. IEEE, 2023, pp. 7025–7030.
- [17] S. Tombaz, S.-w. Han, K. W. Sung, and J. Zander, "Energy efficient network deployment with cell DTX," *IEEE Communications Letters*, vol. 18, no. 6, pp. 977–980, 2014.
- [18] A. Chatzipapas, S. Alouf, and V. Mancuso, "On the minimization of power consumption in base stations using on/off power amplifiers," in *2011 IEEE Online Conference on Green Communications*. IEEE, 2011, pp. 18–23.
- [19] S. K. G. Peesapati, M. Olsson, M. Masoudi, S. Andersson, and C. Cavdar, "Q-learning based radio resource adaptation for improved energy performance of 5g base stations," in *2021 IEEE 32nd Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*. IEEE, 2021, pp. 979–984.
- [20] D. Renga, Z. Umar, and M. Meo, "Trading off delay and energy saving through Advanced Sleep Modes in 5G RANs," *IEEE Transactions on Wireless Communications*, 2023.
- [21] T. Smejkal, J. Bierbaum, T. Oberhauser, H. Schirmeier, and H. Härtig, "Sleep well: Pragmatic analysis of the idle states of intel processors," in *Proceedings of the IEEE/ACM 10th International Conference on Big Data Computing, Applications and Technologies*, 2023, pp. 1–10.
- [22] O-RAN Alliance, "Technical Specification; O-RAN Control, User and Synchronization Plane Specification 11.0; O-RAN.WG4.CUS.0-R003-v11.00," O-RAN Alliance, Technical Report, 3 2023.
- [23] J. Ji, X. Ren, L. Cai, and K. Zhu, "Downlink scheduler for delay guaranteed services using deep reinforcement learning," *IEEE Transactions on Mobile Computing*, vol. 23, no. 4, pp. 3376–3390, 2023.
- [24] C. Marquez, M. Gramaglia, M. Fiore, A. Banchs, C. Ziennicki, and Z. Smoreda, "Not all apps are created equal: Analysis of spatiotemporal heterogeneity in nationwide mobile service usage," in *Proceedings of the 13th International Conference on Emerging Networking Experiments and Technologies*, ser. CoNEXT '17. New York, NY, USA: Association for Computing Machinery, 2017, p. 180–186. [Online]. Available: <https://doi.org/10.1145/3143361.3143369>
- [25] A. Santoro, D. Raposo, D. G. Barrett, M. Malinowski, R. Pascanu, P. Battaglia, and T. Lillicrap, "A simple neural network module for relational reasoning," *Advances in neural information processing systems*, vol. 30, 2017.
- [26] P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Liò, and Y. Bengio, "Graph attention networks," in *International Conference on Learning Representations*, 2018.
- [27] P. J. Huber, "Robust estimation of a location parameter," *Breakthroughs in statistics: Methodology and distribution*, pp. 492–518, 1992.
- [28] J. A. Ayala-Romero, A. Garcia-Saavedra, and X. Costa-Perez, "Risk-aware continuous control with neural contextual bandits," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 38, no. 19, 2024, pp. 20 930–20 938.
- [29] D. Silver, G. Lever, N. Heess, T. Degris, D. Wierstra, and M. Riedmiller, "Deterministic policy gradient algorithms," in *International conference on machine learning*. Pmlr, 2014, pp. 387–395.
- [30] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski *et al.*, "Human-level control through deep reinforcement learning," *nature*, vol. 518, no. 7540, pp. 529–533, 2015.
- [31] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, "Continuous control with deep reinforcement learning," *arXiv preprint arXiv:1509.02971*, 2015.
- [32] A. El-Amine, M. Iturralde, H. A. H. Hassan, and L. Nuaymi, "A distributed Q-Learning approach for adaptive sleep modes in 5G networks," in *2019 IEEE Wireless Communications and Networking Conference (WCNC)*. IEEE, 2019, pp. 1–6.
- [33] F. Salahdine, J. Opadere, Q. Liu, T. Han, N. Zhang, and S. Wu, "A survey on sleep mode techniques for ultra-dense networks in 5G and beyond," *Computer Networks*, vol. 201, p. 108567, 2021.
- [34] J. Kim, H.-W. Lee, and S. Chong, "Traffic-aware energy-saving base station sleeping and clustering in cooperative networks," *IEEE Transactions on Wireless Communications*, vol. 17, no. 2, pp. 1173–1186, 2017.
- [35] J. A. Ayala-Romero, J. J. Alcaraz, A. Zanella, and M. Zorzi, "Online learning for energy saving and interference coordination in hetnets," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 6, pp. 1374–1388, 2019.
- [36] P. Chang and G. Miao, "Optimal operation of base stations with deep sleep and discontinuous transmission," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 11, pp. 11 113–11 126, 2018.
- [37] F. E. Salem, T. Chahed, E. Altman, A. Gati, and Z. Altman, "Optimal policies of advanced sleep modes for energy-efficient 5G networks," in *2019 IEEE 18th international symposium on network computing and applications (NCA)*. IEEE, 2019, pp. 1–7.