The Energy Footprint of 5G Multi-RAT Cellular Architectures

MARTIN KLAPEZ[ID]**, CARLO AUGUSTO GRAZIA**[ID]**, (Member, IEEE),**
AND MAURIZIO CASONI[ID]**, (Senior Member, IEEE)**
Department of Engineering Enzo Ferrari, University of Modena and Reggio Emilia, 41125 Modena, Italy

Corresponding author: Martin Klapez (martin.klapez@unimore.it)

ABSTRACT While environmental and financial reasons point to the minimization of the number of radio sites to contain energy use, emissions, and operating expenses of cellular networks, the industry has, in retrospect, partly headed in the opposite direction. The reasons are the deployment of new technologies alongside the existing ones, the use of higher frequencies, and the expansion of radio coverage. In some cases, the issue has been addressed by replacing legacy access sites with Multi-RAT infrastructures, composed of individual radio elements that run multiple technologies concurrently. Now, the transition to 5G poses additional challenges. This paper reviews Multi-RAT architectures, outlines their benefits, discusses how 5G can be integrated, and provides guidance in terms of architectural recommendations, all from an energy consumption standpoint. Specifically, we firstly summarize the transition of monolithic base stations into modern radio elements, exposing the energy rationale and discussing the impact of each main network component. From this basis, we survey different deployment strategies and evaluate their energy implications in light of the constraints and opportunities given by the 5G New Radio standard, its flagship applications, and its transport requirements. Then, we lay down energy-saving estimates quantifying the contribution of each main network segment, revealing the most promising architectures, and identifying the main challenges and the research directions ahead.

INDEX TERMS Base stations, cellular networks, energy efficiency, mobile communication, network topology, radio access networks.

I. INTRODUCTION

The Radio Access Network (RAN) accounts for the majority of power consumption in the lifetime of a cellular network, including production, transport, usage, and disposal [1], also considering data centers, offices, and stores [2]. The amount of expended energy is mainly determined by the extent of geographical coverage, not by the generated traffic, even though statistical data of traffic load unequivocally indicates that the average network utilization stays low in both temporal and spatial domains [3]. Traditionally, new RAN infrastructures were deployed alongside the existing ones with each new cellular technology roll-out. This deployment strategy allows for the gradual deployment of new technologies, spreading CAPital EXpenditure (CAPEX) over time, and extends the Return on Investment for older infrastructures. However, it also has drawbacks. OPerating EXpenditure (OPEX) increases together with the maintenance needs of a greater number of network elements. Moreover, the overall power consumption rises as more energy is necessary to keep more extensive cellular networks operating. Considering the current grid sources, as well as the combustion-based generators that power Base Stations (BSs) in off-grid areas or during outages, such growth also expands the use of fossil fuels, implicating a rise in energy costs and harmful emissions. Combustibles directly used by big network operators amount between the dozens and the hundreds of millions of liters per year, while overall energy costs go towards a billion EUR/USD per year per operator [2]. In some cases, during the consolidation process of 4G networks, the industry replaced existing 2G and 3G base stations (BSs) with new infrastructures that run 2G, 3G, and 4G from single radio hardware items.

This paradigm shift has been called Multi-Radio Access Technology (M-RAT) [4], and its further generalization to include 5G together with the energy implications are

The associate editor coordinating the review of this manuscript and approving it for publication was Derek Abbott[ID].

discussed in this article. Firstly, we introduce related work, as well as the sources for the data and assumptions presented in the rest of the manuscript (Section II). We start the discussion by giving an overview of M-RAT infrastructures and explore why and how they may be beneficial in terms of overall energy consumption by dissecting RAN infrastructures into their main components and exposing sources of energy inefficiencies, thereby presenting the energy rationale behind the transition of monolithic base stations into modern radio elements (Section III). We then delve into the access protocol stack, present possible split points together with architectural options, and discuss their limitations and their effect on energy consumption, examining how 5G infrastructures can be integrated into M-RAT frameworks (Section IV). From there, we expose challenges and identify promising deployment strategies in terms of energy savings (Section V). We support the findings with power-saving estimates that take into account the architectural options discussed in the paper, considering and quantifying cascade effects among different network segments and visualizing their relative power savings and normalized efficiency ratios (Section VI). To the best of our knowledge, none of the above is currently covered by the literature. We also pinpoint the main challenges that remain to be addressed together with possible research routes and disclose the potential limitations of our work (Section VII).

II. RELATED WORKS AND DATA SOURCES

The acronym Multi-RAT has been used to indicate several different contexts in the literature. For instance, it is often found in works pertaining to the Internet of Things (IoT), such as [5] or [6], where authors intend M-RAT as the concurrent use of several low-power radio access technologies like LoRaWAN and NB-IoT from a single IoT device to improve performance and energy efficiency, the latter by leveraging synergy opportunities between the different radio technologies. In other cases, M-RAT is associated with the use (at times concurrent) of different kinds of wireless standards, such as IEEE 802.11, IEEE 802.16, and cellular [7] or, more recently, is seen in manuscripts that deal with the possible coexistence of different vehicular networks' standards, such as DSRC and C-V2X [8]. Instead, the paradigm studied in this paper regards the co-location of multiple cellular standards into single radio items, a trend pursued by the industry [2] and also studied in academia [4]. The topic received substantially more consideration with the advent of 5G [9]–[11] with particular attention to dual connectivity, given that the feature has been included in the 3GPP 5G standard [12], [13]. Concurrent connectivity to multiple cellular RATs presents its set of issues; for instance, [14] addresses the problems of joint M-RAT assignments and dynamic power allocation, while [15] approaches the feature probabilistically by studying how mobile equipment may decide to access either multiple RATs or a single one. Heterogeneous networks, where macro and small cells coexist in the same geographical area, are also frequently linked to M-RAT deployments [16], [17].

Often, works address performance concerns but not energy consumption directly, which is our main concern here. Indeed, the reported energy consumption figures of radio items vary depending on numerous factors, such as source, technology, cell size, components, manufacturers, configuration, and radiating power. In general, the older the cellular technology, the more energy is required to power a single radio item. For instance, a single 2G GSM BS may require an average power of 3.8 KW for an annual energy amount of around 33.3 MWh [18]. A single 3G BS may require instead 1 KW for an annual consumption of 9 MWh [18], [19]. Values for an LTE BS may be around 0.5 KW, resulting in an annual consumption of 4.5 MWh [20], [21]. However, this generalization does not take into account cell coverage; while an LTE BS has a spectral efficiency that is 30-40 times that of a 2G BS, to offer the same coverage of the latter, an LTE network may even consume more energy [22]. Given the disparity of information in the literature, our first concern has been to build a solid baseline by averaging data and statistics from the academic and industrial sources that follow. To generate the estimates that will be presented in Section VI, we also established a number of efficiency assumptions and enacted several robust constraints to temper uncertainties and guarantee the estimates to revolve around worst-case scenarios. Both the assumptions and the constraints are listed in Section VI. Our sources provided both aggregated data as well as power consumption and efficiency data for specific components or network segments. For the former, [23] presents power needs of 2G and 2.5G BSs on the basis of cell size and delves into the consumption of their main components, while [24] provides energy consumption figures from on-site measurements of 2G and 3G cell towers. In [25], the authors present a study of 2G and 3G energy consumption based on cell size. [26] studies the energy consumption of an LTE network on the basis of the traffic kind and load the network is subject to. [27] delves into the energy characteristics of Cloud-RAN and several of the optimizations we present in Section V, while [28] explores lean design specifically, as well as the effect of its implementations at different levels of aggressiveness on the power consumption of 5G systems (pre-3GPP-standard). From the industry, instead, aggregated data on the energy needs of cell towers can be found in [29], [30]. Information regarding the power consumption of specific components or network segments is reported in the following sources, published by academia, industry, and standardization bodies. Power supplies' energy needs and efficiency curves can be found in [31]–[39]. The power consumption and climate control systems' energy efficiency ratios appear in [31]–[35], [37]–[42]. Information on the power needs of baseband processing, both local and cloud-based, appears in [31]–[39], [43]–[45], while data on the energy needs of radiofrequency converters, that are strictly related with baseband processing units in monolithic BSs, in [32]–[38] and [45]. Figures on the energy required for power amplifiers, often one of the most impacting components in cell towers, as well as data on their energy
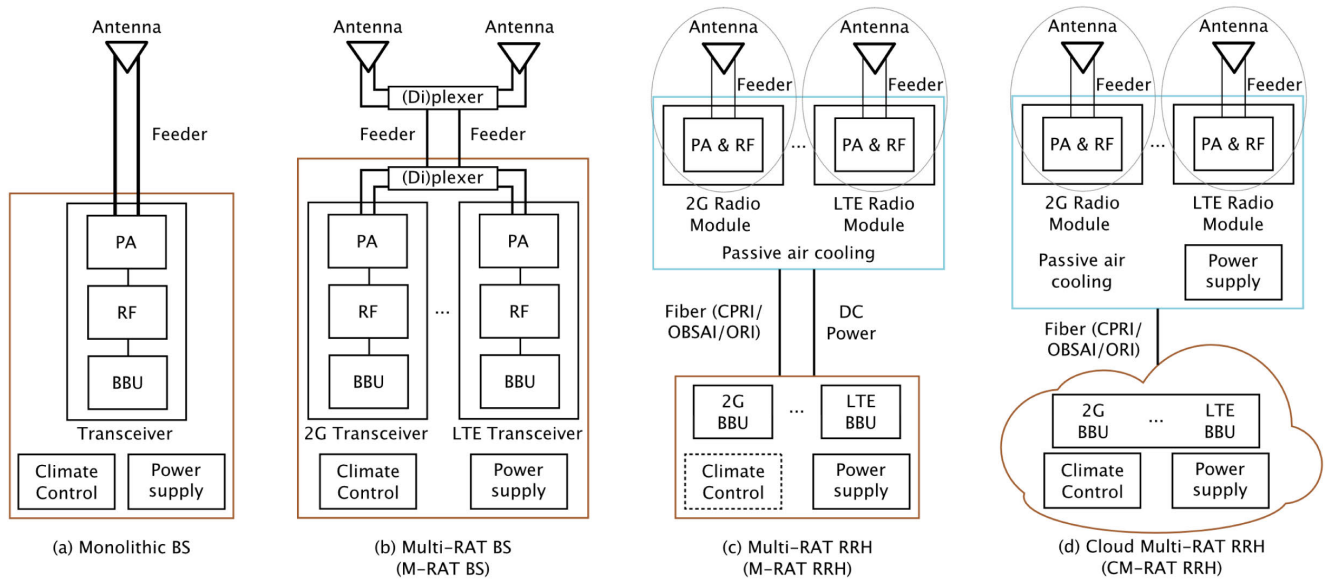
FIGURE 1. Logical components of Monolithic BSs, M-RAT BSs, M-RAT RRHs, and CM-RAT RRHs.

efficiency can be found in [31]–[38], [45]–[49]. In relation to power amplifiers, information regarding feeders in particular is provided in [32], [34], [35], and [37] whilst [32]–[34], [36] explicitly supply data in relation to the output transmission power of a radio item. [32], [37] include considerations pertaining to the relation in power consumption between macro and small cells. Last but not least, information about different fronthauls' energy needs, an often neglected network segment, can be found in [43]–[45] and [50]–[53]. A detailed list of all our sources, tagged for network segment, can be found at [54].

III. M-RAT ARCHITECTURES AND NETWORK SEGMENTS

There are several different methods to co-locate the equipment for M-RATs [55]. If they operate on the same frequency band, frequency refarming allows using a common antenna system, although its azimuths and tilts are also shared. We instead consider the most general case, where each technology uses a different frequency band and, consequently, where a dedicated single or multi-band antenna is employed for each standard. Figure 1 compares the main network components in a traditional monolithic BS, an M-RAT BS, an M-RAT Remote Radio Head (M-RAT RRH), and a Cloud M-RAT RRH (CM-RAT RRH). The presentation considers a single sector for clarity. Brown enclosures comprehend elements typically deployed at ground level, while azure ones those proximal to the antennas. We use the term *network segment* to denote specific sets of components in a network. For instance, the climate control segment is the set of all climate control components in the network.

A monolithic BS is logically composed of a power supply (PSU), a climate control system (CLI), a transceiver, feeder cables (FEEDs), and an antenna. Apart from the latter, each of these elements consumes a significant portion of the energy

used by the BS, and presents opportunities for energy savings. The PSU converts input Alternate Current (AC) into Direct Current (DC) to provide energy to the system. A CLI, usually directly fed by AC, keeps the components' temperature within an appropriate operative range. The transceiver manages input and output signals, and it is in turn mainly composed of three sub-elements: a BaseBand Unit (BBU) that performs digital signal processing, a Radio Frequency (RF) module that converts digital signals into analog signals and vice versa, and a Power Amplifier (PA) that amplifies the signal for transmission through the antenna. The latter is often reported to be one of the main sources of energy consumption in a BS that may reportedly account for more than 50% of its total, especially if considered together with feeders in macro BSs [32], [35], [37]. Typically, the antenna is positioned high above the ground, while the other components are encased in a box at ground or roof level. FEEDs are long coaxial cables that carry RF signals between the ground modules and the antenna. The antenna may be placed very far from its transceiver, producing significant FEED losses that translate into substantial energy waste. Each BS has its backhaul link that connects with the core network.

The M-RAT BS architecture modularizes its transceivers, allowing the coexistence of multiple standards and a certain flexibility in adding or removing features. Each M-RAT BS is composed of a unique PSU, a single CLI, various transceivers, FEEDs, antennas, and a backhaul link that transports data for all the standards implemented in the structure. Operating an M-RAT BS instead of distinct monolithic BSs yields significant energy savings. By powering several transceivers at once, a single modular PSU can, on average, operate with higher efficiency. The same applies to CLI, as a centralized cooling system is typically more efficient than a distributed one. These higher efficiencies alone result in

lower total energy consumption when compared with that necessary for the operation of different monolithic BSs implementing the same standards separately. Furthermore, diplexers such as those shown in Figure 1(b) may cut the total number of FEEDs in half, an improvement that is especially valuable where the distance between ground enclosures and antennas, and, consequently, the relative energy waste, is high.

RRHs are BSs that separate their BBUs from their radio units, usually keeping the former on the ground while enclosing the latter near the antenna. In small cells, both can be found on top. For optimal energy savings, transceiver equipment can be integrated with the antenna to basically nullify FEED losses, saving up to 30% of power alone by significantly reducing one of the main sources of energy waste [2]. Typically, passive air circulation satisfies RRHs cooling needs, although it depends on cell size and environmental temperatures. It is not only hot climates that may result in the need for active air, adiabatic, or liquid cooling systems, as cold environments can, in turn, make heating systems necessary. The presence of an optional CLI is represented in Figure 1(c) with a dashed line.

Figure 1(d) depicts a CM-RAT RRH, which is an M-RAT RRH deployed with the Cloud-RAN (C-RAN) paradigm [56], where the BBUs are pooled and centralized in a remote location. Centralizing BBUs eliminates the need to have power-hungry CLIs in each BS or in each RRH baseband site, which, together with the overall radio site simplification, can reduce OPEX for new sites by ~50%, thanks in a significant part to their lower energy consumption [57]. C-RAN also allows increasing the overall energy efficiency (EE), the computational resources available to signal processing algorithms, the load balancing among cells, and the implementation of multicell-based algorithms. Trials in China with pooled and virtualized BBUs resulted in energy savings of up to 70% compared with legacy BSs [58], i.e., it has been reported an energy expenditure of approximately 1/3 than that required by monolithic BSs offering the same services.

IV. 5G M-RAT ARCHITECTURAL OPTIONS

The overall architecture of the access protocol stack is approximately the same for all cellular technologies, the most prominent difference being the absence of the Packet Data Convergence Protocol (PDCP) from the 2G and 3G stacks (3G includes it but only in the packet-switched user branch). From lowest to highest, the stack is formed by a physical (L1), a data link (L2), and a network layer (L3), as in Figure 2. The first is composed of the RF module and a series of signal processing blocks (PHY). The second is divided into the Media Access Control (MAC), the Radio Link Control (RLC), and the PDCP. The third, which handles Control Plane (CP) as well as User Plane (UP) data, contains only the Radio Resource Control (RRC) (that can be sided by the IP protocol in the UP).
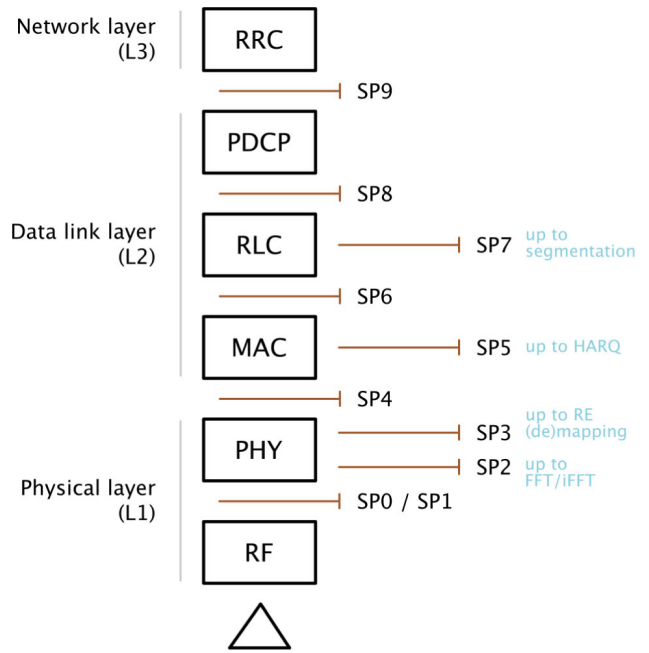
FIGURE 2. General access protocol stack and split points.

A functional split point (SP) separates stack functions among different physical premises. It determines how network functions are aggregated and how the hardware is deployed and, consequently, it can have a significant impact on the energy needs of a network. The C-RAN paradigm positions it between RF and PHY, SP0 in Figure 2, resulting in maximum remote aggregation. However, the CPRI fronthaul interface depicted in Figure 1(c)-(d) was originally designed to transport sampled radio waveforms via short fiber links spanning tens of *m*, not to be stretched to tens of *km* or more. This SP also has other important drawbacks. First of all, the fronthaul consumes energy. This is often overlooked, but while centralizing BBUs can undoubtedly increase their EE, a higher fronthaul capacity generally also means a higher fronthaul power drain. Our estimates in Section VI indicate that it might even need more than the centralized BBUs. Another problem is capacity. The transport of IQ symbols requires a very high continuous bitrate regardless of user traffic; moreover, the bitrate scales linearly with the number of antennas, an enormous burden with Multiple Input Multiple Output (MIMO) systems. The third issue is latency. The fronthaul cannot extend indefinitely because PHY operations require coordination from higher layers that needs to complete in a rigid timeframe. In this regard, the most critical operation is the Hybrid Automatic Repeat Request (HARQ) performed in the MAC, for which the transport is left with only a few hundred *μs* after processing by typical implementations is over [59]. With respect to all, M-RAT deployments complicate things further, as signals of multiple RATs have to be multiplexed on the fronthaul. If 5G is also included in the equation, the load may be unbearable. Even if the fronthaul

could be dimensioned accordingly and the latency could be kept within boundaries, its power consumption would quickly escalate.

Other SPs have been devised to meet demands while keeping some centralization benefits. SP1 follows a compression of time-domain units or a non-linear quantization, reducing fronthaul bandwidth requirements down to 30%-50%. This reduces capacity needs but does not address the scaling problem, creating even more pressure for latency as further processing to execute (de)compression tasks shrinks the already meager transport time window.

SP2 partitions PHY, keeping cyclic prefix and FFT/iFFT local. These operations are load-independent, and the required bandwidth is reduced down to ~1/3 compared with SP0 [57]. However, the bitrate is still both scaling with the antennas and constant, as resource element (de)mapping (that detects unused subcarriers, enabling variable bitrates) is not included, and the fronthaul must be dimensioned and powered accordingly. While SP3 keeps those functions locally, its benefits may be negligible if the fronthaul is dimensioned for maximum utilization, as it must provide for the cases without unused subcarriers and, therefore, it roughly needs the same amount of energy required with SP2. However, there is another option. The instantaneous bandwidth to support depends on the major part on the current load and, therefore, the fronthaul can be deliberately under-dimensioned to only account for non-peak aggregated average loads and yield benefits from load-balancing gains. If the cases of maximum utilization would be rare, this strategy can significantly improve the fronthaul's energy efficiency and power consumption by rationally under-dimensioning it. From SP3 to SP8 included the bitrate scales with the number of MIMO layers instead of the number of antenna ports [59].

SP4 sets the separation between L1 and L2. If dimensioned for maximum capacity, the required data rate is approximately 3% than SP0 and 10% than SP1, SP2, and SP3. The downside is that any joint processing of PHY functions, and as such, the majority of potential energy savings given by BBU aggregation, is ruled out, as higher energy efficiencies may only be extracted by pooling L2 and L3 operations, which represent roughly 20% of the total BBU processing.

Being the HARQ remote, all the former options have very strict transport latency requirements, in the range of μs. From SP5 onwards, the HARQ loop does not cross the split interface, relaxing requirements to the ms scale. This allows fronthauls to span much longer distances. The benefits in terms of energy consumption and bitrate to support, instead, are low. Up to SP7 included, the L2 ARQ is centralized, making those SPs robust over mobility and bad transmission conditions.

The 5G New Radio (NR) interface, as standardized by 3GPP with Release 15 (TR 21.915 V15.0.0), offers two ranges of frequencies, i.e., 0.41 to 7.125 GHz, and 24.25 to 52.60 GHz. BBUs are split into two parts, a Distributed Unit (DU) and a Central Unit (CU), although a Remote/Radio Unit (RU) detached from a DU is often considered as well.

In this paper, we use the RU, DU, CU architecture. The link between the first two is called fronthaul, while the other midhaul. Generally, a DU can be associated with a single CU. A link to a second CU might be put in place but only for backup purposes. A CU, instead, can serve multiple DUs, but the actual topology depends on deployment. The function placement can be adjusted as will on the basis of goals, antenna configurations, and radio channel environmental conditions. 5G infrastructures are being deployed in compliance with the 3GPP Non-Stand Alone (NSA) architecture, where the NR gNodeBs coexist with the LTE eNodeBs and operate under the 4G core. Only 4G services are supported, but they may be delivered with 5G NR access capabilities. The 3GPP Stand-Alone (SA) architecture, instead, allows independent 5G deployments. Both the NSA and SA architectures are compatible with M-RAT deployments, but in the former case, LTE radio items act as master nodes and determine the behavior of the associated NR modules for each user device. The roles can then be exchanged at a later deployment stage. Figure 3 shows two examples of possible 5G M-RAT architectural options. Both are organized into three tiers, i.e., RU, DU, and CU.

Figure 3(a) depicts a double split RAN designed for low-load scenarios, where SP0/SP1 separates the RU from the DU. This architecture may mimic macro cell functionalities. The PSU is optional because DC power can be transferred from the DU to the CU in case they are near. The technologies expected to pressure the fronthaul are LTE and NR, but the split can still be considered practical given the low bandwidth to support. The DU embeds L1 functions only, establishing SP4 as the second split. PHY functionalities are pooled because interference mitigation, and, in general, roughly 80% of the BBU processing happens there. As such, it accounts for the most energy savings that can be extracted from BBU aggregation. As the HARQ loop crosses the midhaul split interface, transport latencies have to stay within a few hundred μs at most. This is the most critical point, that may prevent the architecture viability if Ultra-Reliable Low-Latency Communication (URLLC) applications or heavy spikes of traffic have to be supported. CU functions can be pooled by both technology and similarities/common subroutines. The former aggregation is beneficial because it allows cache-related and preemption mechanisms in hardware processors to shorten processing times, which in turn enables them to spend more time in low-energy states. The latter, instead, can reduce signalling and coordination traffic among elements, leading to further energy savings.

Figure 3(b) shows a double split RAN designed for high-load scenarios. Such architecture can be suitable for small cells. The fronthaul still originates from an SP0/SP1 between the RUs and the DU, but the HARQ loop does not cross the midhaul channel, which, therefore, has relaxed requirements thanks to SP8. This also allows multiple blocks of RUs-DU to be connected to the same CU. Instead, the latency-critical points are the fronthaul links, plural, because multiple RUs can be deployed for frequency reuse. Depending on the
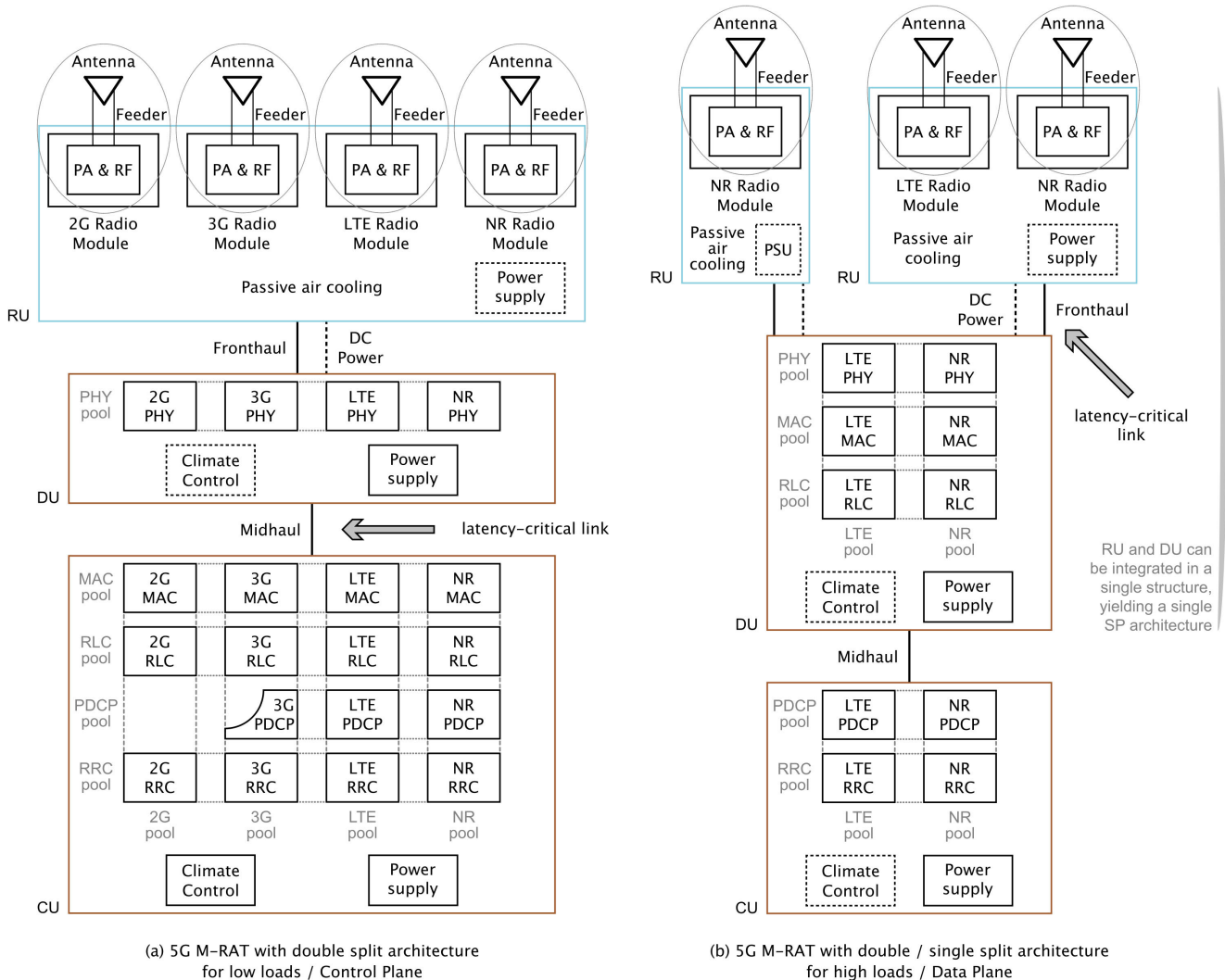
FIGURE 3. Two examples of possible 5G M-RAT deployments; (a) shows a double split, SP0/SP1 for the fronthaul and SP4 for the midhaul; (b) is a SP0/SP1 - SP8 division, but the RUs and the DU can be merged.

scenario, this may leave the door open to URLCC applications, heavy traffic, and aggregation-based energy savings at the same time. If convenient, DUs can be collapsed with RUs, closing the bottleneck and resulting in a single split architecture.[1]

V. DEPLOYMENT STRATEGIES AND OPTIMIZATIONS

SA frameworks will use a service-based core network architecture, where the constituting elements are defined in terms of network functions rather than by hardware entities. RAN as a Service (RANaaS) applies the concept to the access network, is applicable also to NSA architectures, and it would be ideal for reducing the overall energy consumption while being able to guarantee strict performance figures when needed, latency, in particular. The SPs are shifted

dynamically, by actively assigning functionalities among network tiers based on current service requirements [60]. A 5G M-RAT implementation with a RU, DU, CU architecture and a double flexible SP (M-RATaaS) would enable energy savings through maximum function aggregation without committing to a fixed architecture. This is shown in Figure 4. For minimum latency, DUs would operate like in a traditional M-RAT RRH, or M-RAT BS if RUs and DUs are joined together. In other cases, baseband functions can be divided between DUs and CUs. Radio items may interact with a virtual BBU, without caring if the baseband services are provided by a local, middle-tier, remote unit, or a combination of the three. Overall, it can be expected that the more functionalities are run in aggregation premises, the lower the overall energy consumption and the likelihood to power local active CLIs.

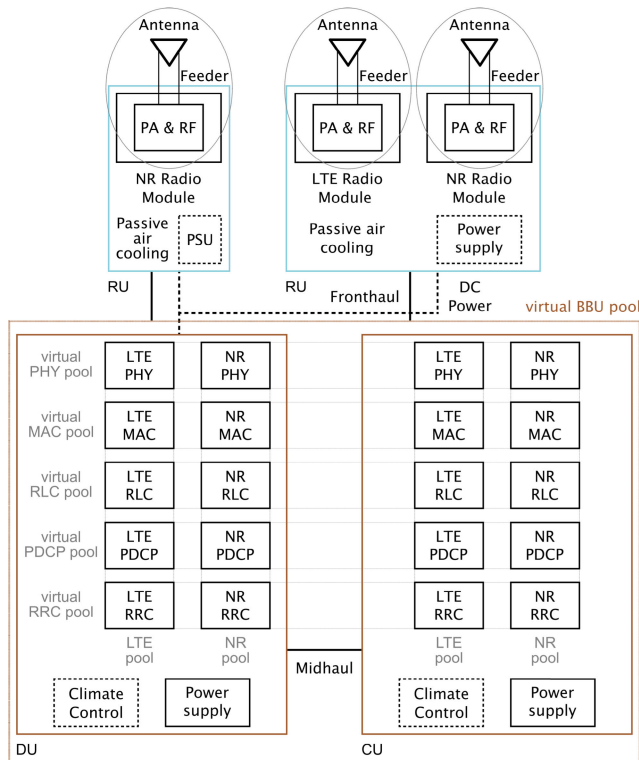M-RATaaS underlines the urgent need to decouple power usage from the sole presence of infrastructures and couple

---

[1]This also applies to the previous architecture, but it generally makes less sense because its latency-critical link is the midhaul.

FIGURE 4. Dynamic double / single split M-RATaaS architecture.

it instead with the actual traffic load. Traditional BSs need to continuously signal their presence and monitor the radio channel to be visible by user terminals and discover them; otherwise, deadlocks, where none can detect the other, would occur. Some sleep procedures have been introduced, for instance, cell wilting and blossoming [61], or power modulation based on statistical data, but they have limited efficacy. The process of network densification may easily lead to more random traffic patterns [62], making effective sleep mechanisms even more desirable. Heterogeneous Networks couple traditional macro cells with one or more layers of small cells under the same coverage area. Phantom cells decouple the CP and the DP and assign those to separate RAN layers, usually to macro and small cells, respectively. 5G M-RAT deployments can implement phantom cells using architectures such as those depicted in Figure 3. This approach introduces more flexibility, inter-cell load balancing, EE, and less overhead. More importantly, it opens to the introduction of effective sleep modes. Signaling and listening operations are still continuously carried out by the CP, but each component of the DP can be put in standby when not needed. The approach is already applicable, even with the 5G NSA architecture where, for instance, LTE macro cells can operate the CP on behalf of higher-frequency NR access nodes. Typically, more than half the energy that would have been expended can be retained [63], and figures can be even higher in areas with dense 5G deployments and heavily discontinuous loads. Moreover, CLIs can be likely moved out of macro cells while,

in the small cells that still need them, they can be put in standby too.

Further optimizations can lead to even higher energy savings. Macro cells can provide DP capabilities under low-load scenarios, as it would be the case if a 2G device connects to the cell in Figure 3(a). Uplink and downlink can also be decoupled and dynamically assigned to different layers/cells. With coordination, this introduces additional load balancing capabilities to the RAN, decreasing congestion, retransmissions and, therefore, power usage. Congestion can be curtailed further by making each layer use a different frequency band. This is also defined in the 5G NSA architecture, as it removes interference between macro and small cells. M-RAT deployments allow modules inside RUs, DUs, and CUs to be selectively and dynamically deactivated, bringing beneficial cascade effects to climate control and power needs. In the spatial domain, switching between MIMO configurations can improve EE, e.g., by providing high data rates with 4×4 instead of 2×2 active antennas, as the former proportionally consume more energy for signals but less for data. In the temporal domain, some subframes of control signals can be neglected to increase transceivers' time in low-power states, while reference signals can be joined with data in ultra-lean design transmissions to minimize signal broadcastings.

VI. ENERGY SAVING ESTIMATES

These estimates assess the relative energy benefits attainable by following the principles and techniques described in the article. Conclusions are drawn based on the relative impact that each network architecture has been found to produce on the power consumption of network components. Those are aggregated in logical network segments, for which totals are reported in the plots. Data has been extracted, extrapolated, computed, and cross-checked from the sources listed in Section II in order to start the analysis with a solid baseline. Following data analysis, the assumptions listed in Subsection VI-A have been devised to provide starting points and to maintain data comparability and clarity. Where applicable, a single SP instead of two has been used to provide conservative figures. We also remark that a significant number of constraints, i.e., those listed in Subsection VI-B, have been enacted to guarantee the estimates revolve around worst-case scenarios.

A. ASSUMPTIONS

The most relevant assumptions regard the average energy efficiency (AEE) of components. The plots report network segments that correspond to components or to the aggregation of components if some are local and others remote. AEE is defined as the ratio between the input and the output power expressed in percentage points. CLI efficiency has been computed starting from the CLI energy efficiency ratio (EER), which we assume to be equal to the coefficient of performance in the case of heat pumps, which will thus be omitted for clarity. The EER indicates the efficiency of a cooling unit in relation to its input power; an EER of 10 means
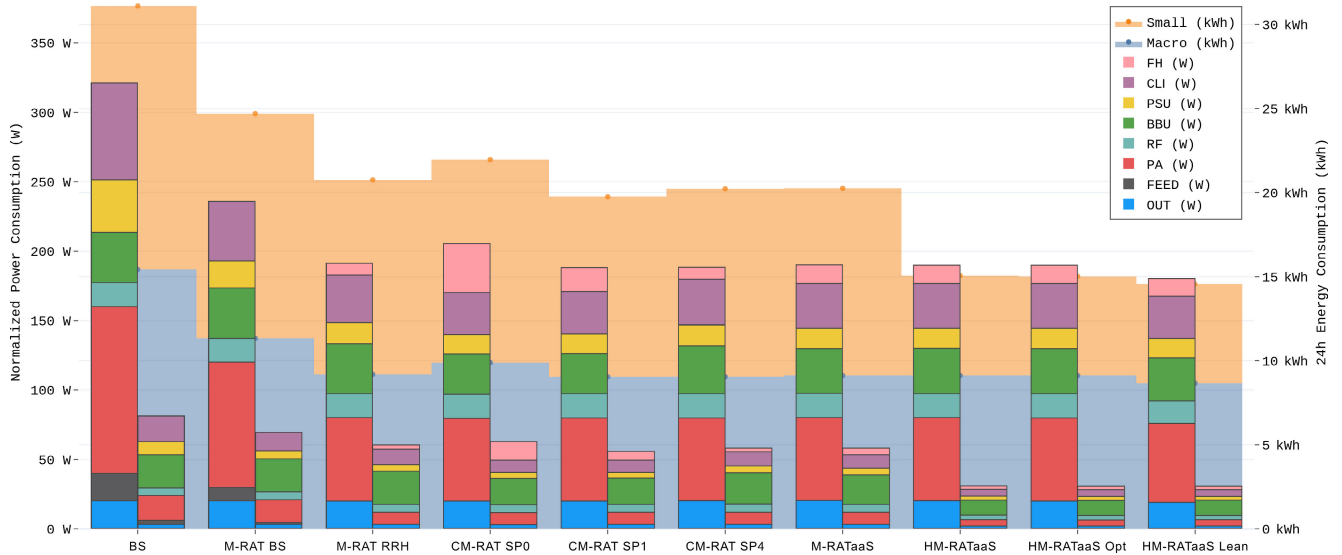
FIGURE 5. Power and energy consumption by logical network segment on different architectures.

the CLI is able to dissipate 10kW of heat by receiving 1kW of input power. We considered the rate at which electronics draw energy as electricity and release it as heat over a time period to be equal, given no physical work is being done and energy is not emitted in any other form such as light. The thermal watts to dissipate have been computed by aggregating the power drawn from all segments except for the CLI itself, output power (OUT), and external FEED. In traditional BSs, PSUs start from an AEE of 85%, FEEDs from an AEE of 50%, and CLIs from an EER of 12. We use 20W and 3W as OUT for each antenna in a macro and small cell, respectively. To keep things comparable and simple, each cell is supposed to be characterized, for each technology, by 3 sectors, 2×2 MIMO, and 2 carriers, for a total of 12 transmission chains. We use 4 technologies per cell, whose power figures are kept aligned for the sake of simplicity. As noted in Section II, in fact, it generally requires more energy to power a cell operating on an older technology than to power a newer one, but, on the other hand, cells based on newer technologies tend to cover a smaller area, at times closing the energy gap if the coverage has to reach previous generations' level. Besides, we are interested in the relative impact that different architectures can have on the energy consumption of each network segment and on assessing the amount of overall energy savings that can be safely assumed to be attainable, not in absolute figures that would inevitably vary on the basis of components, manufacturers, configuration, and deployments.

B. CONSTRAINTS

As estimates are intrinsically imprecise, we enacted the following constraints that act as 'guards' to guarantee the estimates revolve around worst-case scenarios.

- While more efficient PAs exist, we use an AEE of 25% for all.

- Similarly, the efficiency of PSUs has been assumed conservatively, as noted in the previous Subsection.
- In sleep modes, we do not include power modulation techniques.
- All modules in the M-RAT CP are always kept on without accounting for statistically justified sleep intervals.
- Sleep in the DP is activated 50% of the time; in reality, the amount is likely to be higher.
- M-RATaaS is supposed to operate for half the time in minimum-latency configuration and half the time in power-saving configuration. The amount of the latter is likely to be higher in practice.
- While lean design techniques have been reported to reduce consumption up to 35%, we settle for 5% only, as the real-world figure is somewhat unclear.
- In all instances, precision instead of comfort air-conditioning is assumed, thus excluding legacy CLI.
- The reduction of cooling needs in data centers is conservatively set to ~20% only.
- CLIs are always assumed present and active in the CP.
- Although zero active cooling needs are often reported for RF units, we always include them in the CLI requirements.

C. ESTIMATES

Figure 5 shows, for each architecture, the average power consumption of every segment and the overall energy consumption over a 24h period. Taller and shorter box plots depict the power consumed by single macro and small radio items, respectively. Each cell is supposed to be characterized, for each technology, by radio items composed in 3 sectors, 2×2 MIMO, and 2 carriers, for a total of 12 transmission chains. The area charts display the energy consumed in 24h
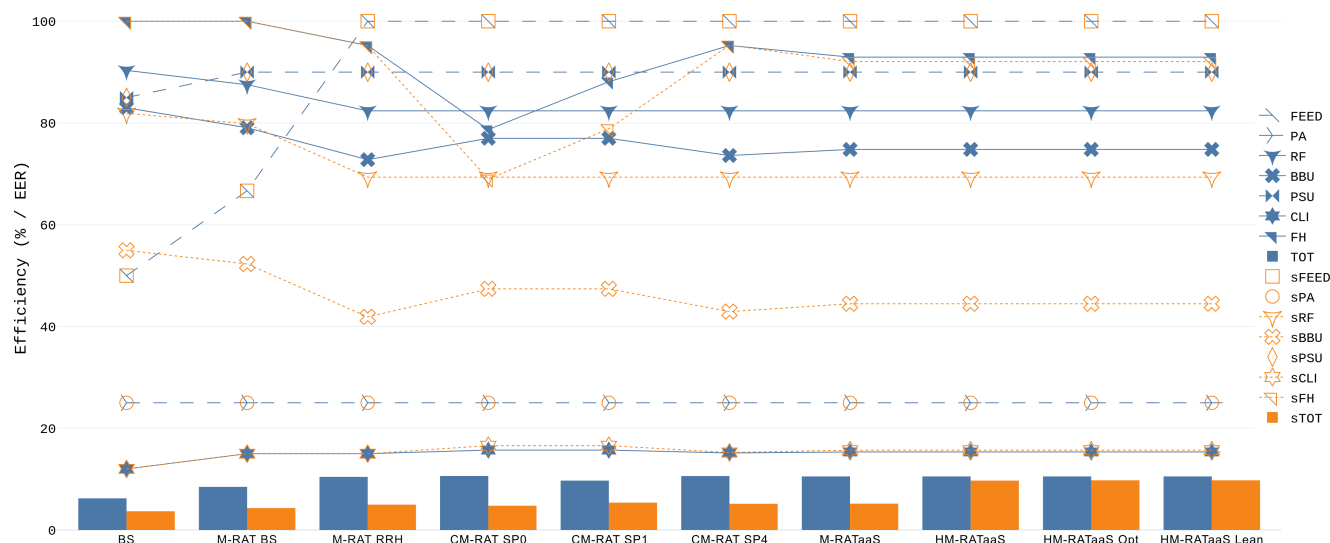
FIGURE 6. Logical network segments efficiency on different architectures.

by a deployment composed of one 2G-3G-LTE-NR macro cell, four 2G-3G-LTE small cells, and four NR small cells. When discussing BSs, they are supposed to be separate cells in the same area. The scatter plots in Figure 6 expose the average efficiency (AE) of each segment in every architecture, defined as the ratio between input and output power expressed in percentage points. The box plots report the total efficiency for macro and small cells, the latter in orange and denoted by a prefixed 's'. For CLI, AE is computed starting from the equipment energy efficiency ratio (EER), which indicates how much W of heat the CLI is able to dissipate for each W of electricity supplied. For RF and BBU, a fixed load is considered. Their AE is inversely computed from the input power of their subsequent segments, i.e., PA and RF, respectively (see Figure 1).

The first thing to point out, not apparent from the plots, is that curtailing the power consumption of a component often reflects positively on other components and so forth, creating a virtuous cascade effect. BSs are mostly affected by high power drains from PAs, CLIs, and, in the case of small cells, BBUs. M-RAT BSs can use half the FEEDs through diplexers, as shown in Figure 1, increasing their AE from 50% to ~67%. Modern PSUs typically have low, maximum, and high efficiency under modest, 50%, and large loads. Using modular PSUs together with the expected higher average power loads given by the co-presence of transceivers for several technologies, the estimated AE is increased from 85% to 90%. Similarly, centralized cooling is normally more efficient, for an estimated EER that shifts from 12 to 15, as shown in Figure 6. In practical terms, the power consumption of PA in macro cells is reduced to approximately 2/3. For small cells, the benefit is less due to lower initial FEED losses. Coupled with their increased AE, the cascade effect on the couple PSU, COOL is significant, with a corresponding reduction of almost 50%, 40% and 40%, 30% for macro and

small cells, respectively. The power needed by RF and BBU does not meaningfully change as their workload remains the same.

Shifting to M-RAT RRHs, the integration of radio equipment inside or very near the antenna systems nullify feeder losses. All the other power contractions are attributable to cascade effects resulting from the FEED removal. Most notably, the PA now needs only half the power that would have required in a BS to produce the same OUT, which in turn further boosts cascade effects on other components. These amount to PSU and COOL reductions around 25%, 20% and 20%, 15% for macro and small cells, respectively. Again, the benefit for the latter is less due to the lower initial FEED. When observing Figure 6, one may think that the weaker AEs of RF and BBU, particularly visible for the sBBU segment, would suggest that M-RAT architectures are worse for RF or BBU-intensive systems. This would be inaccurate as, given their fixed load and the consumption reduction of the other segments, their efficiency appears to decrease because it is inversely computed from the input power of their subsequent segments in the radio chain. There is an actual drawback from the M-RAT RRH onwards, and it is the often neglected power needed by the fronthaul link (FH). This is substantial in the SP0 case, as clear from Figure 5, despite the higher EER of 18 given by data-center-level cooling. The harmful contribution of FH can be particularly daunting in dense small cell deployments, although it depends on the fronthaul topology. The detriment is still significant with SP1 as the channel has to be dimensioned accordingly, but the higher efficiency of the remotely pooled BBU components can compensate in this case. As shown in Figure 6, the BBU segment's efficiency decreases with SP4, as only 20% of the processing is efficiently pooled remotely. Interestingly, the estimates indicate a slightly higher SP4 power consumption overall with respect to SP1, which is to be expected given

the lower BBU aggregation, so far as the transport medium is efficient.

Introducing phantom cells results in another major improvement, as indicated in Figure 5 by the area chart and the small cells bar, mainly due to the DP sleep states. The efficiency of single segments does not change, but the total efficiency of small cells increases, as exposed by Figure 6, due to the lower overall energy spent to provide the same amount of useful OUT. Load balancing optimizations further increase sleep times. In Figure 5, the benefits are not apparent due to the low density of small cells in the considered scenario. For the same reason, the improvements given by lean transmissions are more visible, as they affect the macro cell.

VII. MAIN CHALLENGES, RESEARCH DIRECTIONS, AND LIMITATIONS OF OUR WORK

M-RAT and Transport. M-RAT can exacerbate pressure on transport links. While their increased energy consumption can be offset by centralization benefits, it remains to verify in which cases the additional demands in terms of latency and capacity can be met and till what extent existing systems can be repurposed.

5G HARQ. The subframe length with NR can be reduced from the usual $1ms$ to $31.25\mu s$ [64]. Potentially, this may further shrink the time available for transport. On the other hand, the HARQ process will become asynchronous. What would happen in practice has to be determined. If the HARQ loop is behind a split point, M-RAT deployments with both LTE and NR would need different requirements to be concurrently met, at both transport and processing levels.

Financial Concerns. In converting a network to M-RAT architectures, the most likely approach would be to start clean-slate for new deployments and to upgrade existing ones. Both cases would likely require higher CAPEX than business as usual, yet OPEX would be curtailed. The timeframe to recoup the additional investments is not clear. CAPEX may also be concerning in M-RATaaS deployments with instances of URLLC applications, as RUs or DUs would need to embed the majority if not all the components for baseband processing.

Limitations of our Work. The provided energy-saving figures are estimates designed to assess each architecture's impact. Sources for our data are listed in Section II, assumptions in Subsection VI-A, and constraints in Subsection VI-B. The absolute power values are prudently representative of real equipment but cannot model every variant on the market. The study does not include supported financial considerations. Finally, we may have inadvertently not considered relevant variables. To temper these uncertainties, we carried the analysis under many robust constraints [54], at the risk of over-lowering the benefit estimates.

VIII. CONCLUSION

Overall, M-RAT frameworks consistently result in meaningful energy savings. Eliminating inefficiencies and sensibly aggregating operations produce significant benefits thanks to virtuous cascade effects. While 5G requirements challenge the integration in M-RAT architectures due to increased pressure on transport links, the architectural flexibility of the NR standard might give an opportunity to rethink cellular networks in light of a load-centric philosophy and implement a whole service-based architecture. An infrastructure based on these principles would allow moving past circumstantial solutions, and suggests the possibility to structurally reduce the energy burden without compromising network performance.

REFERENCES

[1] A. Attar, H. Li, and V. C. M. Leung, "Green last mile: How fiber-connected massively distributed antenna systems can save energy," *IEEE Wireless Commun.*, vol. 18, no. 5, pp. 66–74, Oct. 2011.

[2] Vodafone Group. (2019). *Sustainable Business Report 2019, GRI-Compliant Sustainability Report*. [Online]. Available: https://www.vodafone.com/content/dam/vodcom/sustainability/pdfs/sustainablebusiness2019.pdf

[3] I. Chih-Lin, Y. Yuan, J. Huang, S. Ma, C. Cui, and R. Duan, "Rethink fronthaul for soft RAN," *IEEE Commun. Mag.*, vol. 53, no. 9, pp. 82–88, Sep. 2015.

[4] A. Alsohaily and E. S. Sousa, "On the utilization of multi-mode user equipment in multi-radio access technology cellular communication systems," *IEEE Access*, vol. 3, pp. 787–792, 2015.

[5] R. Mozny, M. Stusek, P. Masek, K. Mikhaylov, and J. Hosek, "Unifying multi-radio communication technologies to enable mMTC applications in B5G networks," in *Proc. 2nd 6G Wireless Summit (6G SUMMIT)*, Mar. 2020, pp. 1–5.

[6] K. Mikhaylov, V. Petrov, R. Gupta, M. A. Lema, O. Galinina, S. Andreev, Y. Koucheryavy, M. Valkama, A. Pouttu, and M. Dohler, "Energy efficiency of multi-radio massive machine-type communication (MR-MMTC): Applications, challenges, and solutions," *IEEE Commun. Mag.*, vol. 57, no. 6, pp. 100–106, Jun. 2019.

[7] G. Nagaraja, H. S. Rameshbabu, and Gowrishankar, "A survey of intelligent approach for handoff decision making for long term evolution heterogeneous network," in *Proc. Int. Conf. Smart Syst. Inventive Technol. (ICSSIT)*, Nov. 2019, pp. 1019–1026.

[8] K. Z. Ghafoor, M. Guizani, L. Kong, H. S. Maghdid, and K. F. Jasim, "Enabling efficient coexistence of DSRC and C-V2X in vehicular networks," *IEEE Wireless Commun.*, vol. 27, no. 2, pp. 134–140, Apr. 2020.

[9] D. Klonidis *et al.*, "Int5Gent: An integrated end-to-end system platform for verticals and data plane solutions beyond 5G," in *Proc. Joint Eur. Conf. Netw. Commun. 6G Summit (EuCNC/6G Summit)*, Jun. 2021, pp. 604–609.

[10] L. Huang, J. He, Y.-W. Chen, P. Dai, P.-C. Peng, G.-K. Chang, and X. Chen, "Simple multi-RAT ROF system with 2 × 2 MIMO wireless transmission," *IEEE Photon. Technol. Lett.*, vol. 31, no. 13, pp. 1025–1028, Jul. 1, 2019.

[11] A. Nabil, A. V. Padaki, M. J. Abdel-Rahman, M. ElNainay, A. B. MacKenzie, and J. H. Reed, "On optimal resource allocation in multi-RAT wireless networks with receiver characteristic awareness," *IEEE Trans. Cogn. Commun. Netw.*, vol. 5, no. 1, pp. 103–118, Mar. 2019.

[12] V. Ramaswamy, J. T. Correia, and D. Swain-Walsh, "Modeling and analysis of multi-RAT dual connectivity operations in 5G networks," in *Proc. IEEE 2nd 5G World Forum (5GWF)*, Sep. 2019, pp. 484–489.

[13] V. F. Monteiro, D. A. Sousa, T. F. Maciel, F. R. P. Cavalcanti, C. F. M. E. Silva, and E. B. Rodrigues, "Distributed RRM for 5G multi-RAT multiconnectivity networks," *IEEE Syst. J.*, vol. 13, no. 1, pp. 192–203, Mar. 2019.

[14] A. Alwarafy, B. S. Ciftler, M. Abdallah, and M. Hamdi, "DeepRAT: A DRL-based framework for multi-RAT assignment and power allocation in HetNets," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, Jun. 2021, pp. 1–6.

[15] S. H. Chae, T. Kim, and J.-P. Hong, "Distributed multi-radio access control for decentralized OFDMA multi-RAT wireless networks," *IEEE Commun. Lett.*, vol. 25, no. 4, pp. 1303–1307, Apr. 2021.

[16] G. Ghatak, A. De Domenico, and M. Coupechoux, "Small cell deployment along roads: Coverage analysis and slice-aware RAT selection," *IEEE Trans. Commun.*, vol. 67, no. 8, pp. 5875–5891, Aug. 2019.

[17] M. Peng, Y. Li, Z. Zhao, and C. Wang, "System architecture and key technologies for 5G heterogeneous cloud radio access networks," *IEEE Netw.*, vol. 29, no. 2, pp. 6–14, Mar./Apr. 2015.

[18] T. Chen, Y. Yang, H. Zhang, H. Kim, and K. Horneman, "Network energy saving technologies for green wireless access networks," *IEEE Wireless Commun.*, vol. 18, no. 5, pp. 30–38, Oct. 2011.

[19] S. McLaughlin, P. M. Grant, J. S. Thompson, H. Haas, D. I. Laurenson, C. Khirallah, Y. Hou, and R. Wang, "Techniques for improving cellular radio base station energy efficiency," *IEEE Wireless Commun.*, vol. 18, no. 5, pp. 10–17, Oct. 2011.

[20] G. Auer, V. Giannini, C. Desset, I. Godor, P. Skillermark, M. Olsson, M. A. Imran, D. Sabella, M. J. Gonzalez, O. Blume, and A. Fehske, "How much energy is needed to run a wireless network?" *IEEE Trans. Wireless Commun.*, vol. 18, no. 5, pp. 40–49, Oct. 2011.

[21] C. Han, T. Harrold, S. Armour, I. Krikidis, S. Videv, P. M. Grant, H. Haas, J. S. Thompson, I. Ku, C. X. Wang, T. A. Le, M. R. Nakhai, J. Zhang, and L. Hanzo, "Green radio: Radio techniques to enable energy-efficient wireless networks," *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 46–54, Jun. 2011.

[22] S. Tombaz, A. Vastberg, and J. Zander, "Energy- and cost-efficient ultra-high-capacity wireless access," *IEEE Wireless Commun.*, vol. 18, no. 5, pp. 18–24, Oct. 2011.

[23] V. Sen, J. D. Lal, and S. V. Charhate, "A review: Power consumption at base stations in wireless networks," *Int. J. Eng. Manage. Res.*, vol. 5, no. 1, pp. 260–266, 2015.

[24] C. Lubritto, A. Petraglia, C. Vetromile, F. Caterina, A. D'Onofrio, M. Logorelli, G. Marsico, and S. Curcuruto, "Telecommunication power systems: Energy saving, renewable sources and environmental monitoring," in *Proc. IEEE 30th Int. Telecommun. Energy Conf. (INTELEC)*, Sep. 2008, pp. 1–4.

[25] S. Bhaumik, G. Narlikar, S. Chattopadhyay, and S. Kanugovi, "Breathe to stay cool: Adjusting cell sizes to reduce energy consumption," in *Proc. 1st ACM SIGCOMM Workshop Green Netw. (Green Networking)*. New York, NY, USA: Association for Computing Machinery, 2010, pp. 41–46, doi: 10.1145/1851290.1851300.

[26] M. Yan, C. A. Chan, A. F. Gygax, J. Yan, L. Campbell, A. Nirmalathas, and C. Leckie, "Modeling the total energy consumption of mobile network services and applications," *Energies*, vol. 12, no. 1, p. 184, Jan. 2019. [Online]. Available: https://www.mdpi.com/1996-1073/12/1/184

[27] J. Wu, S. Rangan, and H. Zhang, *Green Communications: Theoretical Fundamentals, Algorithms, and Applications*. Boca Raton, FL, USA: CRC Press, 2016.

[28] A. Prasad, A. Bhamri, and P. Lundén, "Enhanced mobility and energy efficiency in 5G ultra-dense networks with lean carrier design," in *Proc. IEEE Conf. Standards Commun. Netw. (CSCN)*, Oct. 2016, pp. 1–5.

[29] WESTELL Technologies. (2015). *Powering Your Cell Towers, Intelligent Site Management*. [Online]. Available: http://c.eqcdn.com/_74096f19f51aa39ec811dd24c654f021/westell/db/361/31450/file/Powering+Your+Cell+Towers.pdf

[30] Z. Chen and L. Wang, "Green base station solutions and technology," *ZTE Commun.*, vol. 9, no. 1, pp. 58–61, 2011. [Online]. Available: https://www.zte.com.cn/global/about/magazine/zte-communications/2011/1/en_144/224588.html

[31] K. G. Radha. Energy efficiency in cellular networks. Indian Institute of Technology Madras. Accessed: Oct. 25, 2021. [Online]. Available: https://www.naefrontiers.org/44661/Green-Comm-Ganti

[32] G. Auer, O. Blume, V. Giannini, I. Godor, M. A. Imran, Y. Jading, E. Katranaras, M. Olsson, D. Sabella, P. Skillermark, and W. Wajda, "Energy efficiency analysis of the reference systems, areas of improvements and target breakdown," EARTH Project, Tech. Rep. Deliverable D2.3, Jan. 2012.

[33] M. Deruyck, W. Joseph, B. Lannoo, D. Colle, and L. Martens, "Designing energy-efficient wireless access networks: LTE and LTE-advanced," *IEEE Internet Comput.*, vol. 17, no. 5, pp. 39–45, Sep./Oct. 2013.

[34] S. N. Roy, "Energy logic: A road map to reducing energy consumption in telecom munications networks," in *Proc. IEEE 30th Int. Telecommun. Energy Conf. (INTELEC)*, Sep. 2008, pp. 1–9.

[35] L. M. Correia, D. Zeller, O. Blume, D. Ferling, Y. Jading, I. Gódor, G. Auer, and L. V. Der Perre, "Challenges and enabling technologies for energy aware mobile radio networks," *IEEE Commun. Mag.*, vol. 48, no. 11, pp. 66–72, Nov. 2010.

[36] H. K. Boyapati, R. V. R. Kumar, and S. Chakrabarti, "Energy consumption assessment in LTE baseband of eNodeB and guidelines for green baseband subsystems," in *Proc. Annu. IEEE India Conf. (INDICON)*, Dec. 2010, pp. 1–4.

[37] M. Alsharif, J. Kim, and J. Kim, "Green and sustainable cellular base stations: An overview and future research directions," *Energies*, vol. 10, no. 5, p. 587, Apr. 2017. [Online]. Available: https://www.mdpi.com/1996-1073/10/5/587

[38] I. Tunaru, H. Rivano, and F. Valois, "WSN-inspired sleep protocols for heterogeneous LTE networks," in *Proc. 10th ACM Symp. Perform. Eval. Wireless Ad Hoc, Sensor, Ubiquitous Netw.*, 2013, pp. 113–120.

[39] N. Rasmussen. Calculating space and power density requirements for data centers. Schneider Electric. Accessed: Oct. 25, 2021. [Online]. Available: https://download.schneider-electric.com/files?p_Doc_Ref=SPD_NRAN-8FL6LW_EN

[40] MIDEA Group. (2015). *Precision Air Conditioning Room Base Station AC, Data Sheet*. [Online]. Available: http://midea.lt/wp-content/uploads/2017/05/Midea-katalogas-Precision-Air-Conditioning-Room-Base-Station-AC-UAB-Anaga-2015.pdf

[41] STULZ GmbH. (2014). *A Guide to Precision Air Conditioning & Cooling Solutions in Data Centres, Server Rooms and All Critical Environments*. [Online]. Available: https://fdocuments.in/document/stulzcompleteprecisionclimatecontrolrange1012enpdf.html

[42] US Department of Energy and US Environmental Protection Agency. *Air-Source Heat Pumps and Central Air Conditioners Key Product Criteria, Product Standard Specifications*. Accessed: Oct. 25, 2021. [Online]. Available: https://www.energystar.gov/products/heating_cooling/heat_pumps_air_source/key_product_criteria

[43] M. Fiorani, S. Tombaz, J. Mårtensson, B. Skubic, L. Wosinska, and P. Monti, "Modeling energy performance of C-RAN with optical transport in 5G network scenarios," *J. Opt. Commun. Netw.*, vol. 8, no. 11, pp. B21–B34, 2016.

[44] H. Gupta, M. Sharma, A. Franklin, and B. R. Tamma, "Apt-RAN: A flexible split-based 5G RAN to minimize energy consumption and handovers," *IEEE Trans. Netw. Service Manage.*, vol. 17, no. 1, pp. 473–487, Mar. 2020.

[45] H. Wang, "Interplay between capacity and energy consumption in C-RAN transport network design," KTH Roy. Inst. Technol., Stockholm, Sweden, Tech. Rep. TRITA-ICT-EX-2016:127, 2016.

[46] H. Hirata, K. Totani, T. Maehata, T. Shimura, M. Take, Y. Jurokawa, M. Onishi, Y. Ada, and Y. Hirata, "Development of high efficiency amplifier for cellular base stations," *SEI Tech. Rev*, vol. 2010, pp. 47–52, Apr. 2010.

[47] C. Ramella, A. Piacibello, R. Quaglia, V. Camarchia, and M. Pirola, "High efficiency power amplifiers for modern mobile communications: The load-modulation approach," *Electronics*, vol. 6, no. 4, p. 96, Nov. 2017.

[48] Microelectronics Technology. *3G/LTE Remote Radio Head (RRH) G21RRH-48-01B, Product Overview*. Accessed: Oct. 25, 2021. [Online]. Available: https://www.taiwanexcellence.org/en/award/product/25830

[49] Taiwan's Bureau of Foreign Trade and Green Trade Project Office. *3G/LTE Remote Radio Head (RRH) G21RRH-48-01B, Product Sheet*. Accessed: Oct. 25, 2021. [Online]. Available: https://www.greentrade.org.tw/en/cp/award/3glte-remote-radio-head-rrh-g21rrh-48-01b

[50] N. Carapellese, A. Pizzinat, M. Tornatore, P. Chanclou, and S. Gosselin, "An energy consumption comparison of different mobile backhaul and fronthaul optical access architectures," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Sep. 2014, pp. 1–3.

[51] I. A. Alimi, A. L. Teixeira, and P. P. Monteiro, "Toward an efficient C-RAN optical fronthaul for the future networks: A tutorial on technologies, requirements, challenges, and solutions," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 1, pp. 708–769, 1st Quart., 2018.

[52] Q. Liu, T. Han, N. Ansari, and G. Wu, "On designing energy-efficient heterogeneous cloud radio access networks," *IEEE Trans. Green Commun. Netw.*, vol. 2, no. 3, pp. 721–734, Sep. 2018.

[53] D. Brubaker, W. Qian, M. Sussmann, Y. Takafuji, M. Akhter, and T. Hiatt. (Jun. 2016). The emerging need for fronthaul compression. ALTERA Corporation and Intel Corporation. [Online]. Available: https://www.intel.com/content/dam/www/programmable/us/en/pdfs/literature/wp/wp-01265-the-emerging-need-for-fronthaul-compression.pdf

[54] UNIMORE Networking Lab. (Sep. 2021). *Data Sources, Assumptions, and Constraints for the Energy Saving Estimates*. [Online]. Available: http://netlab.unimore.it/mrat_energy_estimates_details.html

[55] J. T. Penttinen, *The LTE-Advanced Deployment Handbook: The Planning Guidelines for the Fourth Generation Networks*. Hoboken, NJ, USA: Wiley, 2016.

[56] A. Checko, H. L. Christiansen, Y. Yan, L. Scolari, G. Kardaras, M. S. Berger, and L. Dittmann, "Cloud RAN for mobile networks—A technology overview," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 1, pp. 405–426, 1st Quart., 2015.

[57] U. Dötsch, M. Doll, H. Mayer, F. Schaich, J. Segel, and P. Sehier, "Quantitative analysis of split base station processing and determination of advantageous architectures for LTE," *Bell Labs Tech. J.*, vol. 18, no. 1, pp. 105–128, Jun. 2013.

[58] I. Chih-Lin, C. Rowell, S. Han, Z. Xu, G. Li, and Z. Pan, "Toward green and soft: A 5G perspective," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 66–73, Feb. 2014.

[59] L. M. P. Larsen, A. Checko, and H. L. Christiansen, "A survey of the functional splits proposed for 5G mobile crosshaul networks," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 1, pp. 146–172, 1st Quart., 2019.

[60] P. Rost, C. J. Bernardos, A. D. Domenico, M. D. Girolamo, M. Lalam, A. Maeder, D. Sabella, and D. Wübben, "Cloud technologies for flexible 5G radio access networks," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 68–76, May 2014.

[61] A. Conte, A. Feki, L. Chiaraviglio, D. Ciullo, M. Meo, and M. A. Marsan, "Cell wilting and blossoming for energy efficiency," *IEEE Wireless Commun.*, vol. 18, no. 5, pp. 50–57, Oct. 2011.

[62] M. Agiwal, A. Roy, and N. Saxena, "Next generation 5G wireless networks: A comprehensive survey," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 3, pp. 1617–1655, 3rd Quart., 2016.

[63] Y. Li, T. Jaing, K. Luo, and S. Mao, "Green heterogeneous cloud radio access networks: Potential Techniques, performance trade-offs, and challenges," *IEEE Commun. Mag.*, vol. 55, no. 11, pp. 33–39, Nov. 2017.

[64] *5G Wireless Fronthaul Requirements in a PON Context*, document ITU-T G.Sup66, Revision 2, Jul. 2019. [Online]. Available: https://www.itu.int/rec/dologin_pub.asp?lang=e&id=T-REC-G.Sup66-202009-I!!PDF-E&type=items

CARLO AUGUSTO GRAZIA (Member, IEEE) received the Ph.D. degree from the Department of Engineering Enzo Ferrari (DIEF), University of Modena and Reggio Emilia (UNIMORE), in 2016. He is currently an Assistant Professor holding the course "Automotive Connectivity." He has been involved in the EU FP7 Projects E-SPONDER and PPDR-TC. His research interests include computer networking, with an emphasis on wireless networks, queueing algorithms, and V2X.

MARTIN KLAPEZ received the Ph.D. degree from the Department of Engineering Enzo Ferrari (DIEF), University of Modena and Reggio Emilia (UNIMORE), in 2017. He is currently a Postdoctoral Research Fellow with UNIMORE. He has collaborated with the Italian nanoscience National Research Center S3, and he has been involved in the EU FP7 Project PPDR-TC. His research interests include verge around network softwarization, public safety networks, and safety-related V2X systems.

MAURIZIO CASONI (Senior Member, IEEE) received the M.S. (Hons.) and Ph.D. degrees in electrical engineering from the University of Bologna, Italy, in 1991 and 1995, respectively. He is currently an Associate Professor in telecommunications with DIEF, University of Modena and Reggio Emilia (UNIMORE), Italy. In 1995, he was with the Computer Science Department, Washington University in St. Louis, MO, USA, as a Research Fellow. He has been responsible at UNIMORE for the EU FP7 Projects E-SPONDER and PPDR-TC.

• • •