IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 12 June 2025, accepted 7 July 2025, date of publication 18 July 2025, date of current version 4 August 2025.

Digital Object Identifier 10.1109/ACCESS.2025.3590365

SURVEY

Energy Efficiency in Network Slicing: Survey and Taxonomy

ADNEI WILLIAN DONATTI[1], MARCIA CRISTINA MACHADO[1], MARVIN ALEXANDER LOPEZ MARTINEZ[2], SABINO ROGÉRIO S. ANTUNES[2], ELI CARLOS FIGUEIREDO SOUZA[2], SAND L. CORRÊA[3], TIAGO C. FERRETO[4], (Member, IEEE), JOSÉ AUGUSTO SURUAGY[2,5], (Life Senior Member, IEEE), JOBERTO S. B. MARTINS[6], (Life Senior Member, IEEE), AND TEREZA CRISTINA CARVALHO[1]

[1]Universidade de São Paulo (USP), São Paulo 05508-090, Brazil
[2]Universidade Federal de Pernambuco (UFPE), Recife 50740-560, Brazil
[3]Universidade Federal de Goiás (UFG), Goiânia 74690-900, Brazil
[4]Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS), Porto Alegre 90619-900, Brazil
[5]CESAR School, Recife 50030-220, Brazil
[6]Universidade Salvador (UNIFACS), Salvador 41770-235, Brazil

Corresponding author: Adnei Willian Donatti (adnei.donatti@usp.br)

This work was supported in part by the Slicing Future Internet Infrastructures (SFI2), Instituto Ânima (IA) under FAPESP/MCTIC/CGI Cooperation Agreement through the Thematic Research Project under Grant 2018/23097-3; and in part by the National Council for Scientific and Technological Development (CNPq).

ABSTRACT Network Slicing (NS) is a fundamental feature of 5G, 6G, and future mobile networks, enabling logically isolated virtual networks over shared infrastructure. As data demand increases and services diversify, ensuring Energy Efficiency (EE) in NS is vital (not only for operational cost savings but also to reduce the Information and Communication Technology (ICT) sector's environmental footprint). This survey addresses the need for a comprehensive and holistic perspective on energy-efficient NS by reviewing and classifying recent strategies across the NS life cycle. Our contributions are threefold: (i) a thorough review of state-of-the-art techniques aimed at reducing energy consumption in NS; (ii) a novel taxonomy that organizes strategies into infrastructure, path/route, and slice operation levels; and (iii) the identification of open challenges and research directions, with a focus on systemic, cross-layer, and AI-driven approaches. By consolidating insights from recent developments, our work bridges existing gaps in the literature, offering a structured foundation for researchers and practitioners to design, evaluate, and improve energy-efficient network slicing systems.

INDEX TERMS Artificial intelligence, energy-efficient slicing, energy-efficient slicing strategy, energy efficiency, network slicing, taxonomy.

I. INTRODUCTION

Energy consumption and its environmental consequences are pressing concerns for modern society, particularly in the design of computer and telecommunication systems. In quantitative terms, the Communication Technology (CT) sector was responsible for an estimated 1.8% to 3.9% of global Greenhouse Gas (GHG) emissions in 2020 [1]. A more alarming perspective from Andrae [2] and Andrae and Edler [3] suggests that this figure may have been as high as 6.3% in 2020, with projections indicating a potential increase to 23% by 2030.

While the use of renewable energy sources offers a promising path to reducing environmental impact, challenges related to their supply, distribution, and scalability remain significant [2]. Consequently, given the current global energy generation matrix (still heavily reliant on fossil fuels), energy consumption from computer systems is expected

The associate editor coordinating the review of this manuscript and approving it for publication was Rentao Gu.

to continue contributing substantially to environmental degradation [4].

In this context, Network Slicing emerges as a promising paradigm that not only addresses performance and scalability demands but also opens opportunities for energy efficiency. By enabling resource sharing and dynamic allocation, NS can significantly contribute to reducing overall energy consumption.

Network Slicing is a technological enabler that allows the sharing of physical infrastructure among multiple virtualized network instances, promoting high resource utilization and operational efficiency. Widely adopted in computer systems and modern telecommunication networks (especially in 5G and beyond deployments), NS plays a pivotal role in improving energy optimization. Furthermore, it is expected to be a cornerstone in future 6G networks, which demand advanced virtualization, adaptive flexibility, and intelligent resource orchestration [5].

Achieving high EE in Network Slicing means addressing a global environmental and sustainability concern and, no less important, reducing costs for the computer and telecommunication sectors.

In this context, several energy-efficient methods, algorithms, and strategies for computer systems and telecommunications have been studied. However, a gap exists in addressing energy efficiency in NS. Motivated by this, this work surveys and proposes a taxonomy to contribute to the community's efforts to promote EE in Network Slicing.

This study aims to answer the following question: "How can one contribute to increasing energy efficiency in network slicing?". To answer this question, the paper identifies the state-of-the-art of energy efficiency in NS and proposes a taxonomy to classify the strategy and methods identified.

The main contributions of this study can be summarized as follows:

- A thorough review of state-of-the-art techniques aimed at reducing energy consumption in NS;
- A novel taxonomy that organizes strategies into infrastructure, path/route, and slice operation levels; and
- The identification of open challenges and research directions, with a focus on systemic, cross-layer, and AI-driven approaches.

The sections of this paper are structured as follows. Section II presents the related work. Section III provides the reader with some background, reviewing NS fundamental concepts and associating them with the energy efficiency scenario. Section IV identifies the strategies used to increase EE in NS. Section V presents the new taxonomy for EE methods in NS and maps the surveyed work on this taxonomy. Section VI discusses the taxonomy and possible research directions and, finally, Section VII presents the final considerations.

II. RELATED WORK

Although many papers have been written regarding EE in 5G networks, fewer of them consider NS. In fact, up to the submission of this paper, the classification of methods and techniques for energy efficiency in network slicing is often addressed in a generalist fashion by research topics or technology, for example, resource allocation, network planning, and energy harvesting. However, we want to provide a taxonomy based on the adopted strategies for achieving energy efficiency gains, classifying them according to the optimization level in which they apply: infrastructure, path/route, and slice operation. Existing works [6], [7], [8], [9] are relevant to our proposal and are further presented in Table 1. We compare them to our work according to the survey's target, the review of EE techniques, the proposal of a taxonomy including specific aspects of NS with EE method, fundamentals review on NS, and the 3rd Generation Partnership Project (3GPP) NS life cycle.

The authors in [6] cover EE techniques in ultra-dense Heterogeneous Networks (HetNets). The overview of EE techniques lists five categories: **1.** Network Planning and Deployment; **2.** Optimization of Radio Transmission Process; **3.** Base Station Sleeping Strategy; **4.** Hardware Solution; and **5.** Energy Harvesting and Transfer. The document structure presents every category with techniques and their respective publications. The authors also discuss the problem formulation of EE in HetNets, presenting results in the literature for power consumption models and energy efficiency metrics, just for a Base Station (BS) without any comment about functional splitting nor NS. The work ends by discussing future directions and lessons learned, contributing to summarizing the scenario. Our proposal is similar to that of [6], as we want to investigate EE techniques for NS, which can be applied to HetNets. Even though it presents recent works on EE techniques in wireless mobile networks, and some physical techniques such as sleep mode and Radio Frequency (RF) transmission which are also included in our taxonomy, their contributions lack investigation on the NS technique, a cornerstone for future networks.

The survey presented in [7] was published in 2016 when 5G was still in its beginnings. The authors propose a taxonomy for energy-efficient techniques for 5G networks grouping them into four broad categories: **1.** Resource Allocation; **2.** Deployment & Planning; **3.** Hardware Solutions; and **4.** Energy Harvesting & Transfer. Back then, the authors pointed out the need for a holistic approach combining multiple techniques. They included in hardware solutions the cloud and edge based implementation of a Radio Access Network (RAN) with Network Function Virtualization (NFV), but without any mention about NS. In our research, the evolution in this field is clear. We have identified energy-efficient techniques based on slice orchestration, considering an end-to-end (E2E) view of all deployed slices. Furthermore, given the evolution of research in the context of NS in 5G, we propose a novel taxonomy classifying techniques according to their strategy to increase energy efficiency. We also summarize and represent techniques according to the optimization level they achieve: **1.** Slice

TABLE 1. Related work comparison.

Ref.	Survey Target	EE Techniques Review	Taxonomy — NS-specific EE Techniques	NS Review	NS lifecycle oriented
[6]	HetNets	✓	✗	✗	✗
[7]	5G Networks	✓	✗	✗	✗
[8]	5G and Beyond RAN	✓	✗	✗	✗
[9]	Softwarized Networks	✓	✗	✓	✗
This work	**E2E Network Slicing**	✓	✓	✓	✓

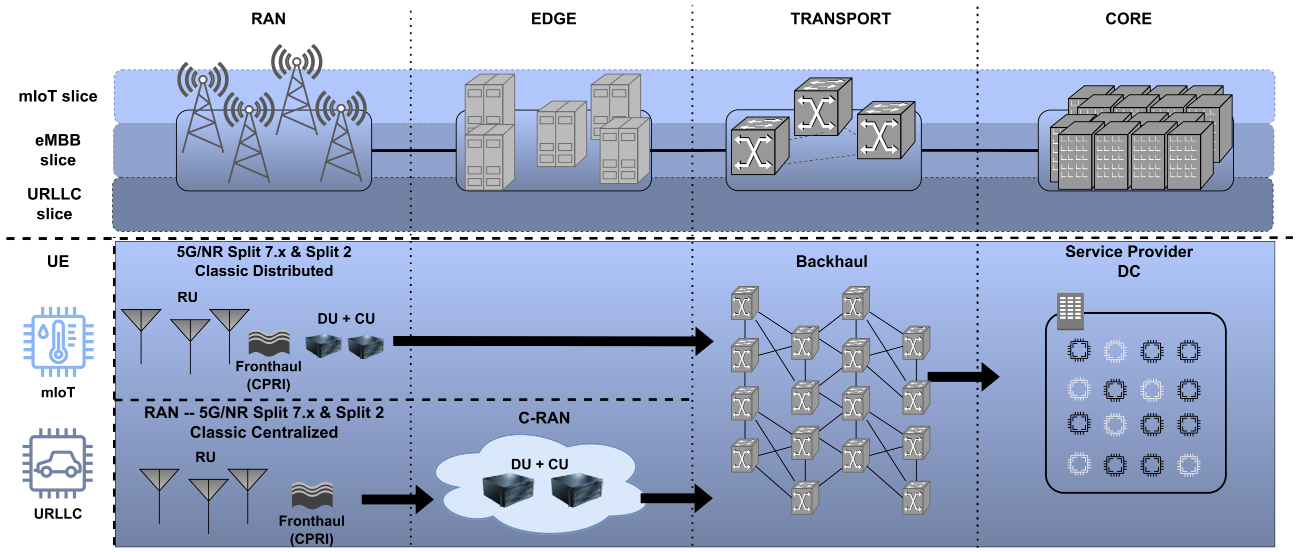

FIGURE 1. Network Slicing in a 5G scenario.

Usage Optimization; **2.** Path/Route Optimization; and **3.** Infrastructure Optimization.

In [8], the authors present a study on energy efficiency for 5G/6G RAN. Their research focuses on RAN network architecture, technology evolution, and network sharing. In the network architecture topic, they discuss the importance of the Virtualized Network Functions (VNFs) location of processing versus its impact on performance parameters such as delay in task execution and data transmission. Considering energy efficiency in RAN, they discuss three main architectural approaches: Virtual RAN (vRAN), Cloud-RAN (C-RAN), or Open RAN (O-RAN). vRAN supposes that VNFs can be run in any hardware and software components that are more energy efficient. C-RAN is a more centralized option. Energy efficiency is achieved by concentrating processing in a central location, which can result in a loss of network performance. Finally, O-RAN is based on open interfaces in mobile networks, allowing the adoption of multi-vendor solutions that consume the least amount of energy. In the case of technological evolution, the main technologies involve improvements in the power amplifier, spectral efficiency, reduced signaling, sleep modes, network virtualization, and Artificial Intelligence (AI).a

They also present several energy improvements enabled by 6G, such as zero touch networks, rate splitting, intelligent reflecting surfaces, improved sampling techniques, and energy harvesting. In the sequence, they briefly present methods for achieving CO2 neutrality.

Finally, they discuss network sharing, which includes: roaming; sharing the entire RAN (except radio frequencies) between multiple operators; sharing of the RAN as a whole (including radio frequencies) among multiple core networks; and NS. At the end, the paper provides guidelines for an operator to move towards greener networks and a map of the different RAN energy-savings possibilities versus consumer experience impact, considering three levels of difficulty in implementing these approaches (minimum, medium, and maximum effort).

The survey carried out by Setiawan et al. [9] covers EE in Software-Defined Network (SDN), NFV, and NS, classifying works according to their network segment (that is, data center, transport, wireless, and emerging: IoT, satellite, vehicular, etc.), evaluation method, metrics (*e.g.*, capacity, latency), and layer. The authors fit EE strategies from the three different subjects (SDN, NFV, and NS) into six categories: 1. Hardware-based improvements; 2. Dynamic Adaptation (DA); 3. Sleep Modes (SM); 4. Heterogeneous Network (HT); 5. Energy Harvesting (EH); and 6. Machine Learning (ML). Sharing these categories among SDN, NFV, and NS results in a broad general classification, giving up details that are explored by each EE mechanism according to the specific subject. For example, approaches that rely on Dynamic Adaptation, Machine Learning, or Sleep Mode may work differently depending on if they are applied to NS or SDN. They conclude their paper discussing future research challenges that range from programmable hardware to experimental evaluation environments. For each softwarized technology (including NS) they classify the works according to their network segment, while we aim to provide a taxonomy considering the specific aspects and details of NS strategies according to their optimization level.

Unlike previous works that offer high-level taxonomies or general discussions of EE across mobile networks, our proposed taxonomy presents a more focused and detailed classification centered on Network Slicing (NS). It organizes EE strategies based on multiple dimensions, including optimization level (infrastructure, path/route, and slice operation) and the underlying mechanisms employed. Additionally, our taxonomy incorporates both the network segment and the slice instance life cycle, providing alignment with end-to-end management frameworks as defined in 3GPP-compliant NS deployments. By doing so, our work offers a clearer and more targeted understanding of how EE is approached specifically within network slicing (a perspective not thoroughly explored in prior studies).

III. BACKGROUND

A. NETWORK SLICING FUNDAMENTALS

The concept of *Network Slicing* was introduced in 2015 by the Next-Generation Mobile Networks (NGMN) Alliance [10], [11]. According to [12], a network slice is "*a logical network that provides specific network capabilities and network characteristics, supporting various service properties for network slice customers*". The network slicing technology is based on "slicing" the physical network into several virtual networks. Each of these virtual networks has its own requirements (*e.g.*, latency, throughput, security) granted by employing VNF chains. In addition, the combination of SDN and VNFs makes it possible to achieve better flexibility and management of the virtualized network and services, benefiting from network programmability, flow forwarding, lower cost, elasticity, and load balancing [13], [14]. Furthermore, these virtualization techniques allow network operators to move from specialized hardware and software to cloud-based solutions. For example, in 5G mobile networks, the RAN segment can be split so that some functions previously associated to the BS are now virtualized and hosted in a C-RAN. Figure 1 illustrates three slices with different requirements, massive IoT (mIoT), enhanced Mobile Broadband (eMBB), and Ultra-Reliable and Low-Latency Communications (URLLC), sharing the same underlying physical infrastructure. This figure also shows how the Distributed Unit (DU) and Central Unit (CU) functions (traditionally associated to the Base station) become virtual functions inside C-RAN.

Most Network Service Providers (NSPs) or Internet Service Providers (ISPs) take on the role of Slice Providers (SPs), since these operators already own the infrastructure. In this sense, SPs can also provide Network Slice-as-a-Service (NSaaS) depending on the customer demand. Organizations such as ETSI-3GPP and 5GPPP play an important role in standardization, providing definitions and recommendations for service providers [12], [15]. In this sense, the slice provisioning workflow is separated into three phases, the network slice life cycle (Figure 2) [12]:

- **Preparation**: During this phase, the network slice is prepared for instantiation. This phase includes tasks such as slice design, capacity planning, and network function evaluation.
- **Commissioning**: This phase deals with the provisioning of the network slice instance. The main tasks that take place in this phase are resource allocation and configuration to meet Quality of Service (QoS) requirements, which can lead to the creation or modification of existing instances;
- **Operation**: This phase begins with the activation of the slice instance. The operation phase also includes tasks related to supervision, monitoring of Key Performance Indicators (KPIs), capacity planning, modification, and deactivation of slice instances; and
- **Decommissioning**: This phase handles tasks such as releasing previously allocated resources, removing configurations, and deleting an instance.

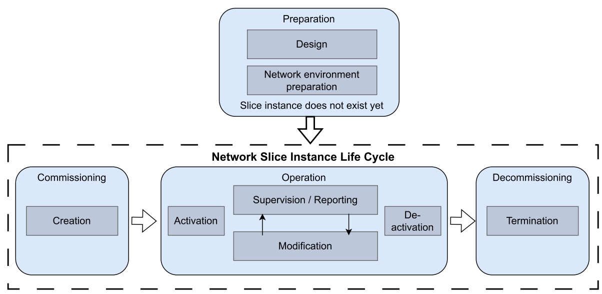

FIGURE 2. Network slicing instance life cycle.

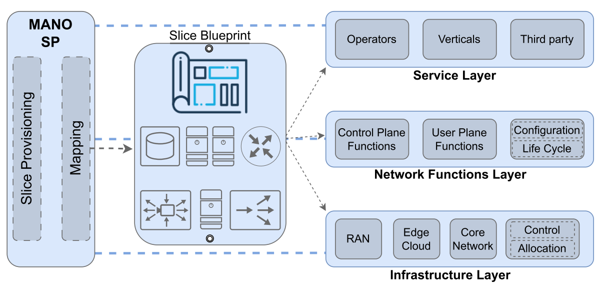

FIGURE 3. Generic Framework representing basic elements and layers in Network Slicing (adapted from [16]).

According to [16], the architectural vision of NS applied to 5G can be mapped onto a generic framework composed of infrastructure, network functions, and service layers.

In addition, a Management and Orchestration (MANO) entity manages and coordinates the elements inside each layer (Figure 3).

In the framework depicted in Figure 3, the infrastructure layer comprises the physical network infrastructure (e.g., routers, cables, switches), including its deployment, control, and management (e.g., allocation of physical resources). The Network Functions Layer is concerned with the configuration and life cycle management of VNFs, which are placed on top of the virtual infrastructure created and chained together to offer an E2E service that meets all the specified requirements. Finally, the Service Layer, which distinguishes Network Slicing from other forms of slicing that have existed in the past (mainly in testbeds and cloud environments), comprises the E2E service description that will be mapped onto a blueprint of all resources and VNFs needed for the network slice deployment. Traditionally, the service description reflects all the business requirements through an Service Level Agreement (SLA) (e.g., availability, latency, and throughput) [16].

B. ENERGY EFFICIENCY AND MACHINE LEARNING IN NETWORK SLICING

NS can potentially reshape the conventional deployment, orchestration, and operation of traditional services, even in global communication network initiatives. However, as noted in [17] and [18], the complexity of the scenario pressures for solutions based on Machine Learning (ML) and, even broadly, AI. In this sense, AI becomes an essential tool in the slicing process due to its inherent ability to provide optimization and acquire knowledge [19], enhancing tasks such as slice orchestration [20].

Allied with the ability to increase performance optimization, ML can also address energy efficiency optimization. Research into the use of ML for Network Slicing [17], [18] shows that a common strategy that seeks better energy efficiency uses Federated Learning (FL). Centralized ML models force raw data to be transmitted across the network, which consumes more energy than spreading data processing to local computations [21]. In addition, some works tend to optimize energy use by applying ML to reduce radio overhead [22], [23] and CPU allocation [24], [25].

Some of the current general trends in terms of ML-based slicing include:

- Embedded ML agents in the SP architecture [26];
- Distributed learning methods such as FL to provide architectural slicing capabilities [27]; and
- Architectural data-driven approaches that support machine learning methods [20].

While AI/ML techniques enable dynamic slice optimization [17], their deployment introduces complex energy trade-offs that must be carefully balanced. The training phase, for instance, has a cost. Deep Reinforcement Learning (DRL) models for slice orchestration increase energy consumption significantly during learning periods due to gradient computations and parameter updates. Even well-trained models, achieving 23-30% energy reduction in RAN slicing through predictive resource allocation, show a break-even point typically occurring after a couple of weeks of operation [24], [28]. While FL reduces this overhead 10× through distributed model training [29], [30], it still requires periodic global aggregation that consumes 35% additional energy compared to static policies [27].

Some trade-offs extend to architectural considerations in the AI/ML models development. For instance, approaches relying on edge-based inference save energy by avoiding cloud transmissions [23], but require specialized hardware accelerators whose manufacturing carbon footprint may offset operational savings [1]. Moreover, even the ML model chosen has different energy proportionalities. The authors in [31] make an energy and cost analysis comparing different approaches. Simpler ARIMA models, for example, might reduce the energy cost by trading off final accuracy (around 15% lower accuracy). Finally, hybrid approaches combining heuristics for routine decisions and ML for complex scenarios demonstrate significant efficiency while reducing computational overhead [29], [31], [32].

C. ENERGY-EFFICIENCY, DECARBONIZATION AND SUSTAINABILITY

In our web-based, user-oriented, service-oriented, and fully connected society, energy efficiency, decarbonization, and sustainability are issues of great concern. The society concern results from the direct or indirect impact that energy consumption can have on GHG emissions that cause climate change [33]. According to [1], the overall global contribution of ICT emissions is between 2.1 to 3.9% and is concentrated in data centers, networks, and user devices. Networks are responsible for 27% of total emissions.

Network slicing, as an important enabler for most new systems such 5G/6G, Internet of Things (IoT), Smart Grid, and others, contributes significantly to energy consumption worldwide. As a figure of merit, it is estimated that 5G energy consumption will account for approximately two-thirds of 1,5% of global energy consumption. Looking at the energy consumption of 5G mobile networks, the RAN consumes 73% of the energy, network interconnections consume 13%, datacenters consume 9%, and other operations account for 5% of energy consumption [8]. These figures reveal the importance of savings in mobile networks and, above all, in the NS as the main enabler of this technology.

Regarding network slicing, there are distinct pathways to net-zero greenhouse emissions through energy efficiency techniques (Figure 4). These pathways include enforcing energy efficiency in the processing phases of NS and prioritizing service providers, which use renewable energy.

We try to shed some light on these alternatives by surveying and discussing how to reduce energy consumption in the network slicing process, and complementing this discussion, we clarify how energy efficiency, decarbonization, and sustainability interrelate with the network slicing process.

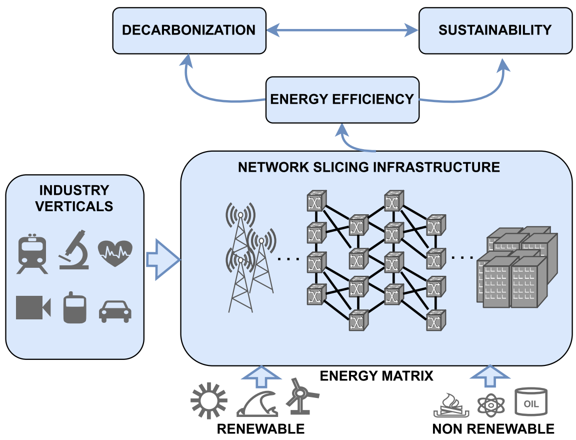

FIGURE 4. Correlation between NS, energy efficiency, decarbonization, and sustainability.

In the computer science domain, energy efficiency refers to the practice and methods of reducing energy consumption to perform a specific task or obtain a particular service. An energy-efficient system aims to maximize the outcomes of the tasks and services for a given amount of energy input, minimizing energy waste and guaranteeing user performance requirements (SLA, QoS, Quality of Experience (QoE), and others) [34], [35]. Anser et al. [36] extend SLA models to slices including energy-based commitments. They also propose including EE metrics in the Virtualized Network Function Descriptor (VNFD).

Decarbonization refers to the ability to reduce or eliminate GHG emissions, such as carbon dioxide (CO2) emissions, towards a net-zero pathway and is mainly associated with human activities. Decarbonization can be achieved through pathways deployed across various economic actors with a predominant focus on the industrial sector, where computer systems are important players [37].

According to the United Nations 2030 Agenda [38] for Sustainable Development, sustainability is a broad concept that involves three main interrelated dimensions: environmental, social, and economic. In line with this, sustainable development must ensure organizational performance while protecting people and the planet [39]. Alternatively, sustainable development can be conceptualized as the ability to meet the needs of the current generation without compromising the ability of future generations to meet their own needs [40].

Environmental sustainability refers to preserving and improving the use of natural resources with strategies that help to ensure that human activity does not compromise the Earth's resources and preserve our ecosystems. Social sustainability works to improve the well-being of the planet's citizens, addressing various issues such as equity, human rights, digital divide, and cultural diversity, among others [39]. Economic sustainability is related to income generation with a consequent improvement in society's quality of life. This can be achieved through technological innovation and new business models, as well as the rational use of resources.

Economic sustainability is linked to network slicing mainly by its ability to reduce carbon emissions by optimizing and reducing energy consumption in the various phases and components involved in the network slicing process. Energy-efficient network slicing also contributes to environmental sustainability with solutions and strategies in pollution control, transport systems, and production systems. When NS-based 5G and 6G technologies are used to support these systems, the contribution can be direct or indirect.

The interrelation of the concepts related to network slicing can be summarized as follows:

- Decarbonization and energy efficiency are closely related concepts that establish a path toward sustainability and promote sustainable energy practices.
- Energy-efficient approaches to network slicing promote decarbonization and lead to a more sustainable use of energy worldwide.
- Another key issue for decarbonization is the deployment of clean and renewable energy sources. Countries have very different energy matrices. Currently, there is an international consensus that all countries must migrate to increasingly renewable and clean energy matrices.

There is a balance in the energy matrix between polluting and renewable, efficient and non-polluting energy generation (Figure 4). Although there are global commitments to increase the use of non-polluting energy, the use of polluting energy will continue to exist. In this way, energy-efficient methods for network slicing will always play an essential role, considering the two scenarios with the use of polluting and non-polluting energy, as they imply a reduction in resource consumption and a reduction in costs for the providers and users of slices.

According to [1], the exact share of renewable energy used in the ICT sector is unknown. Nevertheless, the sector is a significant purchaser of renewable energy, which points to a global shift towards the use of renewable energy use. In global terms, the use of renewable energy could significantly reduce the contribution of ICT's carbon footprint, leading to around 80% reduction in carbon emissions.

Energy efficiency technologies, such as energy management, smart manufacturing, material efficiency, and energy efficiency methods, are well-established pathways and can potentially reduce energy use and GHG emissions [33]. Energy efficiency is a common practice in computer science systems like IoT, ICTs, widely used in smart cities, data centers, and initiatives such as the smart grid, which focuses on the efficient production, transmission, and distribution of energy (Figure 4) [41], [42] [43], [44].

To summarize, we believe that Network Slicing (NS) can contribute to all three pillars in sustainability:

- **Environmental**: By enabling virtualized, on-demand services, NS reduces the need for always-on dedicated infrastructure, directly lowering energy usage. Efficient orchestration can reduce idle resources and facilitate

green scheduling (e.g., sleep mode in underutilized RAN elements) [4], [8];
- **Economic**: NS allows service providers to dynamically instantiate slices tailored to specific verticals, improving Operating expenses (OPEX) through infrastructure sharing [45];
- **Social**: Tailored slices can prioritize underserved areas (*e.g.*, rural 5G FWA) or critical services (e.g., healthcare), contributing to digital equity and social impact [46], [47]

IV. OPTIMIZATION LEVELS TO INCREASE ENERGY EFFICIENCY IN NETWORK SLICING

To identify the main contributions related to Energy Efficiency (EE) in Network Slicing (NS), we defined two search queries and applied them to the IEEE and ACM Digital Libraries, covering the period from 2019 to 2024: "Network Slicing AND Energy Efficien*" and "Network Slicing AND Orchestration AND Energy Efficien*". From the 189 papers initially retrieved, our research team filtered and selected 36 papers for in-depth discussion based on a preliminary screening of titles and abstracts. To systematically analyze the approaches adopted across these works, we developed the taxonomy illustrated in Figure 5, which categorizes EE methods in NS according to their optimization scope. This classification emerged from recurring patterns identified during our discussions, in which the strategies to enhance EE consistently aligned with one of three distinct levels: *Infrastructure Optimization*, *Path/Route Optimization*, and *Slice Operation Optimization*. Each level reflects a different scope of decision-making within NS environments. For instance, Infrastructure Optimization includes methods such as server placement near end-users (e.g., edge computing, BS distribution), which reduce energy consumption through architectural design and resource positioning. We also observed that strategies within the same level can differ in nature. For example, while both cell distribution and spectrum efficiency techniques fall under Infrastructure Optimization, the former focuses on spatial resource allocation, whereas the latter targets technological enhancements. Thus, our taxonomy not only organizes existing contributions but also clarifies the conceptual basis behind EE strategies in NS, offering a structured lens through which researchers and practitioners can evaluate and compare different approaches.

Moreover, the classification into three optimization levels emerged from our in-depth analysis and discussion of the selected literature. Throughout the review process, our goal was to unravel the diverse approaches to increasing EE within the context of NS. By systematically examining the objectives, assumptions, and methodologies of the analyzed works, we observed that EE strategies consistently aligned with one of three distinct optimization scopes. This three-level taxonomy effectively captures the hierarchical nature of energy-related decisions in NS systems, ranging from structural configurations to dynamic operational controls.

Figure 5 represents the methods operating at each optimization level in the pyramid. At the base (**Infrastructure Optimization**), methods increase energy efficiency by optimizing the infrastructure through resource positioning (e.g., functional split, cell distribution) and technology improvement strategies (e.g., sleep mode, power amplifier). By the middle of the pyramid (**Path/Route Optimization**), techniques address energy efficiency by optimizing the slice's path/route. Techniques acting at this level increase energy efficiency by sharing, managing, and optimally allocating resources (resource sharing strategy). At the top of the pyramid (**Slice Operation Optimization**), methods work on slice usage control. Techniques at this level save energy by minimizing or avoiding energy waste in the slice operation context.

A. INFRASTRUCTURE OPTIMIZATION LEVEL

In our research, we observe that methods operating at the infrastructure optimization level focus on resource positioning and technology improvements. The architecture design of a Slice Provider (SP) must consider energy efficiency, which, besides being environmentally friendly, also translates to OPEX reduction. Therefore, we consider infrastructure optimizations as the primary source of energy savings in Network Slicing (NS). In this sense, the overall energy consumption is reduced by planning the computational resource capacity, positioning, and technology adopted in the slice provider infrastructure. For instance, the computational resource capacity should be thought to increase the effectiveness of turned-on devices. Moreover, the adopted technology in the slice provider should also corroborate lowering the energy consumption (e.g., IEEE 802.3az, Energy Efficient Ethernet).

The SP relies on a distributed system to provide network slices. Therefore, the architecture design might benefit from strategies that usually derive from classical distributed systems, which precede the NS concept and are out of the scope of this document. For example, the publications [48], [49], [50], [51] present contributions that are not directly related to the NS energy efficiency (data center/cloud solutions and green Ethernet, respectively). However, it is noteworthy that such classical resource positioning and technology improvements could be extended to the SP infrastructure to reduce energy consumption.

B. PATH/ROUTE OPTIMIZATION LEVEL

The Path/Route Optimization Level is based on the slice orchestration idea. Since the term "orchestration" is broadly utilized in several contexts, we refer to slice orchestration as the coordination and control of the overall workflow process of tasks related to the network slice life cycle (slice preparation, commissioning, operation, and decommissioning). Therefore, we observe how documents address energy efficiency in coordinating and integrating multiple network slices. The slice provider must orchestrate all the network

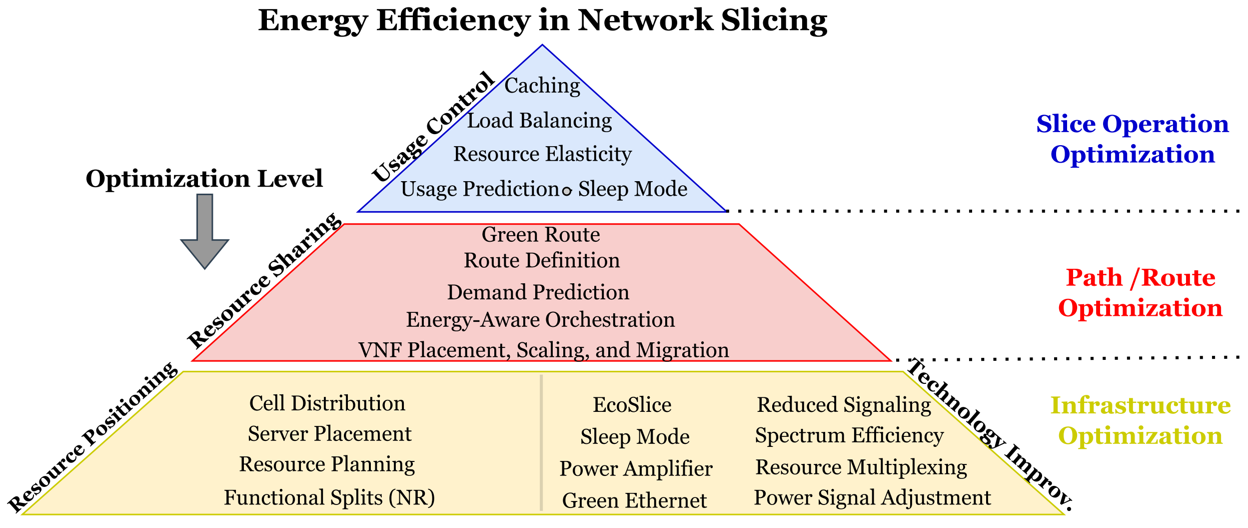

FIGURE 5. Optimization levels to increase energy efficiency in Network Slicing.

slices served by its infrastructure, for example, sharing resources, calculating routes, and predicting traffic peaks in routes. Therefore, techniques operating at this level consider a panoramic view of all the slices in the system (mostly as a network graph).

C. SLICE OPERATION OPTIMIZATION

Methods operating in this optimization level assume that the network slice is already in its operational phase. While approaches based on path/route take the provider's viewpoint (global vision of all slices), methods based on operation optimization take the slice's viewpoint. Therefore, most techniques optimize the usage of computational resources (e.g., computing, network, storage) and workloads.

In Figure 5, the strategy to increase energy efficiency using a sleep mode technique might be based on an infrastructure decision or on slice operation optimization. In case it is a native function from the device (*e.g.*, switch, router), then we consider it a technological improvement, fitting as an infrastructure decision. On the other hand, if the sleep mode is asserted through an algorithm analyzing a container or a server workload, for example, the technique suits operation optimization. Generally, works using strategies based on slice operation optimization benefit from workload and traffic fluctuation.

V. METHODS FOR ENERGY EFFICIENCY IN NETWORK SLICING: A TAXONOMY

In this section, we expand the pyramid presented in Figure 5 to a comprehensive taxonomy of methods for energy efficiency in NS. While the pyramid provides the reader with a layer-based and concise notion of such techniques structuring, we present the surveyed work alongside a detailed taxonomy depicted in Figure 6. Additionally, it is noteworthy that papers in this research field might not address energy efficiency as a specific goal or measured metric. Instead, such documents usually address efficiency in terms of cost or resource usage, for example. Therefore, depending on the context, the gains in efficiency might translate directly into power usage (*e.g.*, CPU and network usage). In the following subsections, we detail surveyed work suiting every optimization level.

A. INFRASTRUCTURE OPTIMIZATION

As mentioned before, Network Slicing can benefit from traditional distributed systems Energy Efficiency (EE) techniques. Additionally, 5G NS brings new challenges and opportunities for conventional infrastructures (mostly in the RAN segment). In this sense, we analyze works in the literature achieving energy efficiency through infrastructure-level optimizations. Efforts in this category take advantage of the following strategies:

- **Resource Positioning Strategy**: resource arrangement favors energy efficiency and overall performance. For instance, choosing the RAN configuration, server placement, and functional splits;
- **Technology Improvement Strategy**: improving or adjusting technologies to reduce energy consumption. For instance, power signal adjustment, resource multiplexing, and Green Ethernet.

The following tables share a common structure and summarize relevant research works that served as references to build our proposed taxonomy and its classification. Each table organizes these studies according to specific network segments, optimization levels, phases defined by the 3GPP standard, and the applied techniques, clearly classified into Artificial Intelligence (AI)-based methods, heuristic methods, and other alternative techniques.

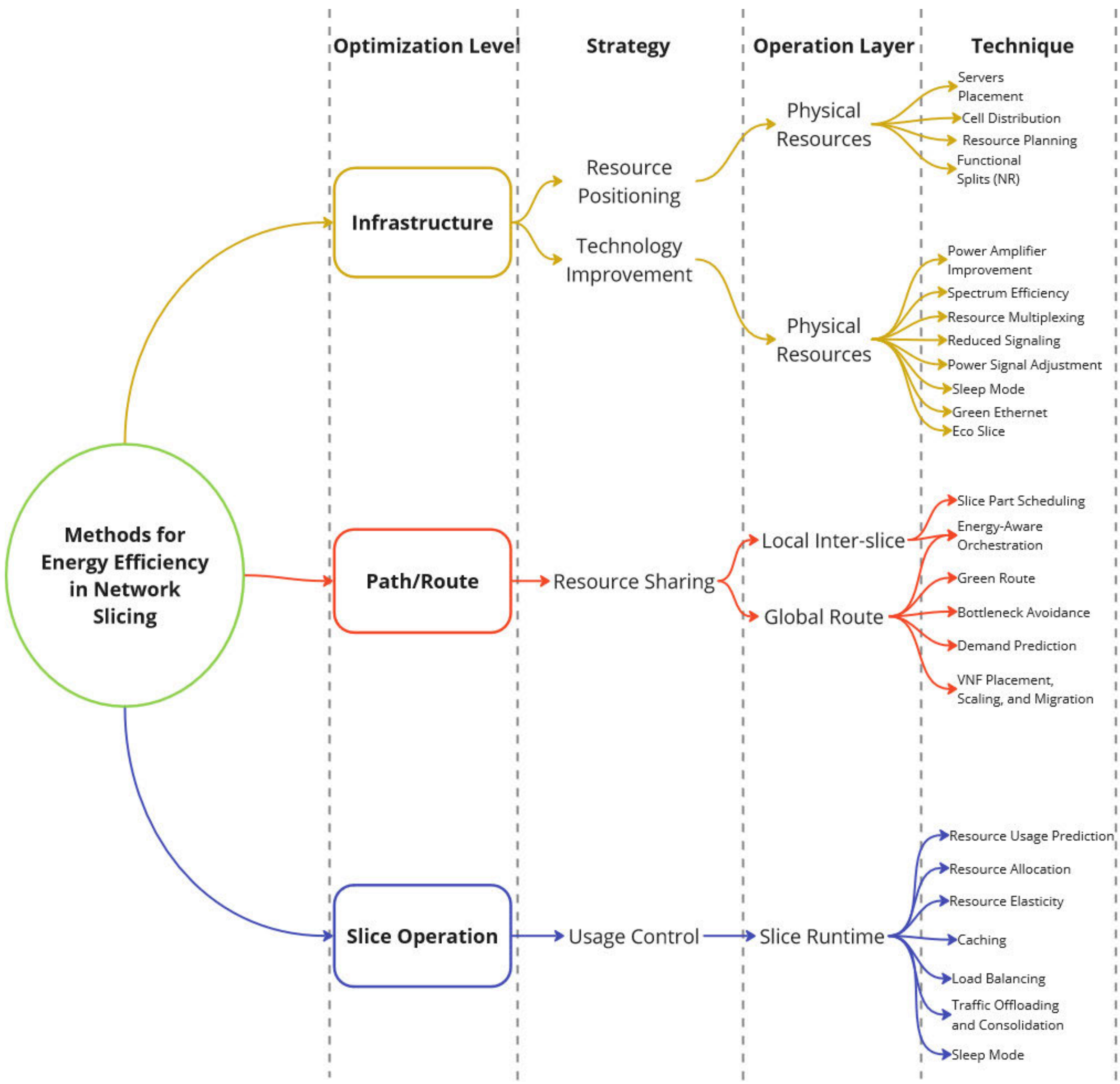

FIGURE 6. Taxonomy of Energy Efficiency Methods in Network Slicing.

Table 2 summarizes works exploring such opportunities in infrastructure-level optimization.

In the work [52], the authors explore energy efficiency in NS by providing optimizations at the infrastructure level, taking advantage of Fog-RAN (F-RAN) to alleviate end-to-end operation (positioning strategy). The authors investigate the impact of fog computing on power consumption and delay performance and propose a joint optimization solution of mode selection (C-RAN mode, fog radio access point mode, and device-to-device mode) and resource allocation in uplink F-RANs for power minimization, considering predefined scenarios. To minimize transmission costs in fronthaul and save system power, authors suggest that local data processing be enabled by Fog Radio Access Points (F-APs) and Fog-UEs (F-UEs). The authors use two Reinforcement Learning (RL) algorithms to maintain low power consumption and stable transmission delay for User Equipments (UEs) and F-UEs.

According to the authors of [53], guaranteeing the slice requirements requires additional energy consumption. In this sense, the authors suggest a dynamic slice activation and deactivation. However, the SP architecture must contemplate a shared slice, called *EcoSlice*, which operates 24/7 and provides bare-minimum service. Thus, in conditions such as low traffic demand, operators may switch users from other slices to the *EcoSlice*. To solve the decision problem of switching users, the authors use two Multi-Armed Bandit (MAB) agents at each BS. The authors claim their solution provides approximately an 11-14% EE improvement. We consider the method presented by the authors as an infrastructure-level optimization since the SP architecture must contemplate a static ''eco-aware slice'' beforehand.

TABLE 2. Research work achieving EE through infrastructure-level optimizations.

Ref.	Network Segment	3GPP Phase	AI Method	Other Methods	Technique
[52]	F-RAN	Commissioning	RL	—	Resource Planning & Positioning
[53]	RAN	Operation	MAB	—	Eco Slice, Sleep Mode
[54]	RAN	Commissioning	—	ILP	Functional Splits (NR)
[55]	RAN	Operation	—	LR	Resource Planning
[56]	RAN, Edge	Operation	—	SACRA	Spectrum Efficiency
[57]	RAN, Edge	Commissioning	—	MINLP	Spectrum Efficiency
[58]	RAN	Operation	—	—	Resource Planning
[59]	RAN	Operation	—	Dinkelbach, others	Spectrum Efficiency
[60]	RMEC	Operation	—	—	Technology Improvement
[61]	RAN, Edge Cloud	Preparation	RL	SCA	Functional Splits, Renewable Energy Integration

In Sen and A.F.A. [54], the energy efficiency is addressed by the RAN functional split, where the function placement strategy considers different functional splits and network slice requirements. This is achieved through an Integer Linear Programming (ILP) heuristics to minimize the energy consumption of processing nodes and transport links considering the slice requirements.

In [59], the authors maximize the energy efficiency of eMBB and URLLC slices through the optimization of remote radio unit selection and beamforming. The problem is formulated as a mixed-integer non-linear programming (MINLP), and simulation results show that the method effectively improves EE while assuring QoS requirements (latency, reliability, and throughput).

Kao and Wu [55] contend that controllability of QoE in 5G networks is a persistent issue. In response, it proposes an AI-enabled architecture designed to predict and sustain QoE. This architecture leverages network slicing and Multi-access Edge Computing (MEC) technologies to gather cross-layer performance data in real-time and dynamically allocate network resources. Although the paper highlights the importance for telecommunications companies to adopt energy-efficient architectures not only to reduce power consumption but also to decrease the carbon footprint, it lacks detailed explanations on how the proposed architecture achieves these reductions. The paper briefly mentions that the architecture could reduce energy consumption by downscaling the number of the allocated Physical Resource Blocks (PRBs), yet this aspect is not extensively explored.

The Non-Orthogonal Multiple Access (NOMA) technique is essential to increase spectrum efficiency in 5G cellular systems. In this sense, Hossain and Ansari [56] propose a technique for integrating NS and NOMA in mobile edge computing, achieving an improvement in spectral efficiency on 5G networks. They propose the Slicing Aware Clustering and Resource Allocation (SACRA) algorithm creating user clustering, computing and wireless resource allocations. Moreover, Ismail et al. [58] review and classify recent advancements in resource allocation for 5G networks, focusing on Power-Domain NOMA (PD-NOMA) techniques. They provide a detailed overview of the existing research into two main categories: power/energy-efficient and rate-optimal resource allocation. In terms of energy efficiency, the paper emphasizes PD-NOMA's potential to enhance power efficiency by allowing multiple users to share the same resource blocks (*e.g.*, time and frequency) while being separated in power levels. This method differs from traditional Orthogonal Multiple Access (OMA) techniques, which allocate separate resources to each user, leading to underutilization and increased power consumption. The paper explains how PD-NOMA can make 5G networks more sustainable and use less energy by giving more power to different users and making it easier to serve many users at once.

Hossain and Ansari [57] investigate the impact of numerology on a sliced Time Division Duplex RAN. The authors also optimize duplex ratio and power and bandwidth allocation. Moreover, they observe that, among other conclusions, the highest numerology schemes do not necessarily translate to the highest average spectrum efficiency.

The authors in [60] present a framework applied to Radio Multi-access Edge Computing (RMEC). In this context, network devices are energy-constrained and powered by batteries. The authors propose automatically generating network slices with wireless energy transfer technology to realize synchronous energy and information transmission by slices. The authors validate their proposed framework by analyzing data from the China Telecom Sichuan Branch.

The document [61] proposes a sustainable and energy-efficient O-RAN-based architecture for deploying 5G Fixed Wireless Access (FWA) in rural areas, where fiber deployment is costly. The authors introduce a three-level closed-loop system that dynamically optimizes radio resource allocation (for O-RAN slices and Customer Premises Equipment) while minimizing energy costs. The edge cloud is powered by renewable energy, and the optimization process uses RL and Successive Convex Approximation (SCA) to balance communication utility and energy efficiency. We consider this work into the Infrastructure Optimization category, as it primarily focuses on network infrastructure improvements (leveraging O-RAN functional splits, renewable energy integration, and intelligent resource management to improve energy efficiency).

B. PATH/ROUTE OPTIMIZATION

Works actuating on the route optimization level observe network slices from a global system perspective. The network graph is configured (and often reconfigured) to save energy. The strategy behind approaches at this optimization level is to distribute and share resources among different network slices sharing the same requirements. For instance, VNF chains can be created based on this idea. Furthermore, demand prediction can also be used to establish more efficient network routes. We summarize publications achieving EE in this optimization level in Table 3.

The work presented in [62] investigates the use of service function chain (SFC) in network slicing. The authors propose minimizing the total power consumption (considering the whole network) by optimally selecting cloud nodes to deploy functions in the SFC. The problem is formulated as a Mixed Binary Linear Program (MBLP) with E2E latency and link capacity constraints. The authors compare their approach with other existing ones. Following a similar logic, in [70], the authors tackle the VNF embedding problem across SFC instances in software-defined 5G networks. They introduce a dynamic VNF sharing model, FlexShare-VNF, which considers flow requests for individual network slices. Implemented as a mixed-integer linear programming (MILP) optimization, the model strategically manages VNF sharing to enhance both network efficiency and hardware utilization. The primary aim is to boost resource utilization through sharing techniques. The authors explicitly link improved resource utilization to more energy-efficient service delivery. Although the paper asserts significant gains in energy efficiency, it primarily quantifies improvements in terms of resource utilization and latency reduction, rather than direct metrics of energy consumption.

The authors of [71] envision 6G networks and its requirements. They propose a novel algorithm running in polynomial time (called TailoredSlice-6G) which efficiently allocates resources within a softwarization context. The algorithm is triggered upon a slice request and then a sub-algorithm is called depending on the required resource type. The authors do not evaluate energy consumption directly. Instead, they rely on the resource consumption ratio. Results show that the proposed algorithm can be used in real networking applications.

Mishra and Hota [63] devise a joint QoS and energy efficiency data-driven approach on top of pre-defined and prioritized traffic classes. Prioritized traffic classes are predicted using Random Forest Regressor (RFR) and Gradient Boosting Regressor (GBR). The solution imposes network dynamic reconfiguration using Cognitive Cycles (CC) to obtain energy savings. Based on simulation results, authors claim node consumption is reduced through the evolutionary CC algorithm, outperforming other standard approaches by at least 23%.

In the paper by Oladejo and Falowo [68], an energy-efficient mixed-integer linear fractional optimization problem (MILFP) is formulated for slice resource allocation considering the Quality of Service (QoS) of different slice use cases. The authors model a global energy efficiency method for the 5G sliced network that provides energy efficiency.

In the publication [65], the authors work on an EE maximization problem considering bandwidth and power optimization. The approach applies to SDN HetNets, by managing radio resources. Furthermore, the authors consider an SDN controller with centralized control and global network information, allowing network slice selection and resource allocation to serve devices.

The publication [66] investigates resource management methods in MEC. The authors design numerical methods optimally allocating edge servers to reduce the average power consumption in the server. Also exploring the benefits of MEC, the work [67] formulates a joint optimization problem of service placement and Small cell Base Station power. The authors aim for a flexible service placement in MEC while saving energy in the base station.

Chen et al. [69] introduce a new framework for the coexistence of unlicensed heterogeneous networks (NR-U and WiFi) utilizing network slicing. Its goal is to implement an efficient network selection and resource allocation strategy to minimize energy consumption while meeting specific user demands for data rate and delay. The authors recommend an online Lyapunov optimization algorithm for optimal scheduling, which is praised for its rapid convergence and low complexity. The approach considers two timescales to mitigate the significant costs associated with energy consumption and performance decline due to frequent network switching. The paper acknowledges a trade-off between energy consumption and traffic delay, which can be managed through adjustable control parameters. For example, the proposed strategy can accommodate a delay-sensitive user by increasing downlink energy consumption to ensure timely transmission, while conserving energy for a delay-tolerant user by extending the transmission duration.

The study [64] presents a framework for managing network resources to improve the performance of the E2E network through the allocation and routing of the VNF with control of delays and capacity restrictions. The study proposes a heuristic that considers traffic from the request point to the user, maximizing the energy efficiency of the virtualized network, calculating the shortest paths, and checking whether this path satisfies the conditional rules. The model considers sound provisioning and heterogeneous slice configurations, including different components of the core radio and transport networks. To allocate resources across network slices, the Network Slice Management Function (NSMF) is used in the instantiation phase supported by the Placement Service, and through algorithms, it establishes the resources that will be allocated, calculating optimization and managing the slice life cycle.

TABLE 3. Publications achieving EE at path/route optimization level.

Ref.	Network Segment	3GPP Phase	AI Method	Other Methods	Technique
[62]	Core	Commissioning	—	MBLP	Routing Optimization & VNF Placement
[63]	Edge, Core	Operation	RFR, GBR	Evolutionary CC	Demand Prediction & Slice Part Scheduling
[64]	E2E	Commissioning	—	HERO heuristic	VNF Placement
[65]	RAN	Operation	—	—	Parametric Dinkelbach & Route Selection
[66]	MEC	Operation	KKT conditions	—	Resource Management & Slice Part Scheduling
[67]	MEC	Operation	—	MILP	Service Placement
[68]	Edge, Core	Operation	—	MILFP	Route Optimization
[69]	E2E	Full life cycle	—	Lyapunov optimization	Network Selection & Green Route
[70]	Core	Commissioning	—	MILP	VNF Consolidation
[71]	Core	Commissioning	—	Polynomial time algorithm	Slice Part Scheduling
[72]	Core	Commissioning, Operation	—	Holt-Winters	Demand Prediction
[73]	Core	Commissioning	RL, QL	—	VNF Placement
[29]	E2E	Operation	FL	Constrained Neural Network	Route Optimization & Slice Orchestration
[74]	Core	Commissioning	GA	Custom K8s Scheduler	Slice Orchestration

In Zhou et al. citebib:72, a dynamic network slice energy-efficient deployment approach is presented. The solution uses a traffic prediction algorithm for network slices coupled with a Virtualized Network Function (VNF) placement and scaling strategy to deploy slices proactively. The strategy of the proposed method is based on resource sharing considering VNF placement and VNF traffic prediction. Authors argue that their method (NSD-CECM) reduces energy consumption compared to other methods (HSR-RSV, ESFC, and VPCM).

The document [74] presents an enhanced scheduling approach for network slicing in 5G, leveraging Kubernetes (K8s) and a Genetic Algorithm-based strategy to improve Virtual Network Embedding (VNE). The authors identify inefficiencies in K8s' default scheduler, which deploys slices pod by pod, often leading to wasted resources (local optima) and increased energy consumption due to incomplete deployments. Instead, their proposed scheduler evaluates the feasibility of deploying an entire slice before committing resources, thereby improving slice acceptance ratio, deployment time, and energy efficiency. This strategy directly contributes to Path/Route Optimization in network slicing by optimizing slice placement and resource utilization, ensuring efficient orchestration within the network infrastructure.

The study [73] presents a scheduling mechanism for a VNF orchestration system based on Kubernetes, as a solution for the efficient use of energy in network slices. The proposed solution uses a reinforcement learning algorithm (Q-Learning) to make decisions about the positioning of VNFs in the network slices, to achieve energy savings.

The document [29] proposes a distributed AI/ML-based network management framework aligned with the MonB5G project. By leveraging FL, the framework decentralizes management and orchestration functions to improve scalability and sustainability in 6G networks. The experimental results demonstrate significant energy savings (over 10x compared to centralized approaches) through proactive resource allocation and prediction mechanisms, specifically CPU scaling for Virtual Reality (VR) streaming servers. Moreover, the paper addresses slice orchestration and resource management across domains (RAN, Core, and Edge) to minimize energy consumption. The proposed solution optimizes paths and routes of network traffic while ensuring service quality, particularly in resource-intensive applications like VR video streaming.

C. SLICE OPERATION OPTIMIZATION

Energy Efficiency (EE) can be achieved in NS also by optimizing the slice usage. Works we found in the literature are listed in Table 4. In this sense, the idea is to reduce resource wastage (e.g., idle devices, overallocation). Methods that apply this strategy lean on the scope of the slice itself. Rather than optimizing the overall resource distribution and management, such methods optimize already granted resources (e.g., caching, load balancing, usage prediction, sleep mode). Therefore, methods allocating resources for an optimized route, for example, belong to the *Path/Route Optimization level* (Section V-B).

Thantharate et al. [31] propose a data-driven Deep Neural Network (DNN) method with Transfer Learning (TL), Autoregressive Integrated Moving Average (ARIMA) and Exponential Smoothing (ETS) to reduce energy consumption in 5G and Beyond 5G (B5G) networks. The approach forecasts traffic load and uses the estimated load to evaluate energy efficiency and OPEX savings. The authors claim that the proposed ECO6G method presents advantages concerning

TABLE 4. Publications achieving EE on the slice usage optimization level.

Ref.	Network Segment	3GPP Phase	AI Method	Other Methods	Technique
[31]	RAN	Operation	DNN, TL	ARIMA, ETS	Traffic Load Prediction & Sleep Mode
[28]	RAN	Commissioning	DRL	—	Resource Allocation
[75]	H-CRAN	Commissioning	DRL	Greedy algorithm	Resource Allocation
[76]	RAN (Sidelink Edge)	Operation	DDQN	—	Resource Allocation
[77]	RAN, Edge	Operation	—	Genetic algorithm	User Association Control
[78]	MEC	Operation	—	BCD	Task Offloading
[32]	E2E	Operation	—	Dataset analysis	Workload Control
[79]	Edge	Operation	RL, DRL	—	Resource Allocation & Elasticity
[80]	RAN	Operation	—	EESO algorithm	Resource Allocation
[81]	F-RAN	Operation	—	Problem decomposition	Caching & Resource Allocation
[82]	RAN, Core	Operation	Multi-Agent DDQN	Joint Optimization Framework	Resource Allocation and usage

other OPEX saving methods. In this sense, energy efficiency is achieved through a usage strategy, enabling network management (e.g., putting base stations in sleep mode) according to usage prediction.

The authors in [28] combine Deep Learning (DL) and DRL for decision-making on resource allocation on a large and small time scale, respectively. Different than most reviewed studies, this work considers not only the radio resources but also the power resources to rate-based and resource-based users in the RAN. The proposed method uses traffic patterns to predict power and Resource Blocks (RBs) allocation for each user in the slice. Simulation results show better energy efficiency and overall performance than other methods published in the literature.

Also concerned with radio resource allocation, the authors in [75] approach a multi-objective problem by maintaining QoS and EE in a Heterogeneous Cloud Radio Access Network (H-CRAN) environment. The authors claim to achieve a better QoS provisioning for the slices while preserving energy consumption compared to other baseband solutions.

Miao et al. in [76] propose a vehicular network (V2X) sideline resource allocation (outside the signal coverage of the 5G base station) method that maximizes energy efficiency and meets communications latency constraints using the DRL Double Deep Q-Network (DDQN) algorithm. The communication resources area is allocated using the dedicated V2V spectrum.

The document [82] addresses energy efficiency challenges in 5G networks supporting heterogeneous services, such as eMBB, Machine Type Communication (mMTC), and URLLC. It proposes a joint optimization framework for subchannel assignment and power allocation using a cooperative multi-agent deep reinforcement learning (MA-DDQN) approach. By dynamically optimizing resource allocation, the approach ensures energy-efficient operation of heterogeneous non-orthogonal multiple access (H-NOMA) networks while maintaining QoS requirements. The authors achieve energy efficiency optimization based on a strategy leveraging real-time adjustments and advanced machine learning techniques. Their method dynamically adjusts operations during a deployed slice's life cycle, improving the resource usage during the slice's operation.

Laroui et al. [79] introduce novel resource allocation algorithms for VNFs in edge computing environments. These algorithms, based on ILP, RL, and DRL, focus on optimizing server utilization, placement time, and energy consumption. The paper showcases how these algorithms can efficiently handle different network scales and complexities, demonstrating their scalability and cost-effectiveness in terms of energy use and server resource optimization. This work is particularly relevant in the context of 5G networks and the increasing demand for efficient resource allocation in edge computing scenarios.

The authors in [81] considered dynamic scenarios for F-RANs where the optimization is performed for each time slot, achieving a flexible control both on EE and delay performance by adjusting the radio, caching, computing, and virtualization parameters. They propose a low-complexity EE optimization algorithm for virtual resource allocation, by decomposing into two sub-problems, the allocation of caching and computing resources and the sub-carrier assignment and power allocation. They also adopt a stochastic model to accommodate virtual resources, time-varying channel conditions and random data arrivals in the simulation analysis. Their simulation results show a 30% improvement in EE under the condition that the guaranteed delay threshold is two slots (i.e., 1 ms).

The study [77] investigates power consumption and energy efficiency in a multi-tenancy scenario for NS in HetNets. The user association problem presumes that the network slices were already created and are currently operational. Thus, the authors use a genetic algorithm to allocate resources for admitted users outperforming baseline schemes in terms of fairness, QoS, and power consumption.

In the document [78], the authors investigate how to offload computation tasks to the MEC in an energy-efficient manner. By adopting MEC, mobile devices' energy and computation capacity is improved. Therefore, the authors formulate and solve the problem using the block coordinate descent (BCD) algorithm. The authors provide numerical results to compare the benefits of their approach over classical schemes.

Bolla et al. [32] propose an AI framework for 6G networks to enhance energy efficiency. This framework aims to create a tight relationship between the workload generated by applications in network slices and the energy consumed by the infrastructure. It introduces a high-level architecture and algorithmic roles for effective implementation. The authors argue that this method can result in considerable energy conservation while satisfying application demands, contributing to sustainable 6G networks.

Ravindran et al. [80] propose an algorithm called Energy Efficient System-resource Optimization (EESO) to achieve energy efficiency by optimizing the resources required/allocated by user services per sub-slice in 5G network. The EESO algorithm is responsible for the optimal allocation of system resources for the services of users connected in multi-sub-slice and the energy efficiency for the services assigned to a user connected by multi-sub-slice.

VI. TAXONOMY DISCUSSION & DIRECTIONS

This section presents key discussions and reflections that emerged during the development and evaluation of the proposed taxonomy for energy efficiency in network slicing. Grounded in the initial research question, *"How can one contribute to increasing energy efficiency in network slicing?"*, this discussion critically analyzes methodological choices, identifies practical and theoretical challenges, and explores the broader implications of the results. By examining the limitations of current approaches, the complexity of implementation in real-world scenarios, and the potential of advanced technologies such as AI-driven digital twins, this section aims to deepen the reader's understanding of the taxonomy's contributions and to stimulate future research. Additionally, it broadens the debate to encompass sustainability considerations, highlighting energy-efficient network slicing as a promising yet underexplored path toward environmentally and socially responsible communication infrastructures.

In this work, we set efforts on a comprehensive classification, introducing a clear and structured taxonomy for energy efficiency in network slicing, which facilitates a more in-depth understanding of existing approaches and guides future research. From our review of the literature and the review of methods and approaches, we understand that current techniques suit one of the three pyramid levels (Figure 5): *Infrastructure Optimization*, *Path/Route Optimization*, and *Slice Usage Optimization*. The literature seems to be concerned about the RAN at the infrastructure level, especially with radio spectrum efficiency. Furthermore, F-RAN, MEC, and different functional splits have been well investigated in terms of resource arrangement and allocation. Moreover, none of the papers we found in our review mention infrastructure-level optimizations at the Core Network (CN) segment. Since this segment is mainly responsible for virtualized network functions, energy efficiency at the infrastructure level often falls back to cloud computing or data center infrastructures.

Resource allocation is a recurring theme in EE research related to network slicing. However, we observed that the term encompasses diverse techniques, depending on the type of resource being optimized (for example, function chaining versus RAN or edge resources). This ambiguity extends to other topics as well, such as usage prediction and user admission control. Thus, previous classifications based solely on thematic categories fail to reveal the actual mechanisms through which EE is achieved. Our taxonomy addresses this shortcoming by explicitly exposing the strategies used to improve energy efficiency in NS. We hope this structured approach can serve as both a reference and a foundation for future research in this field.

While the proposed taxonomy provides a structured understanding of energy-efficient strategies in Network Slicing (NS), it is crucial to acknowledge the practical challenges that network operators face when attempting to implement these solutions in real-world scenarios. First, the integration of EE strategies often requires significant architectural changes, especially at the *Infrastructure Optimization Level*, where methods such as edge server deployment, green hardware adoption, or cell densification demand both financial investment and logistical planning. Furthermore, many techniques assume a high degree of control and observability over the network, which may not be feasible for operators constrained by legacy systems or vendor-specific hardware limitations.

At the *Path/Route Optimization Level*, the orchestration mechanisms that enable energy-aware resource sharing and traffic management rely heavily on accurate, real-time monitoring and centralized control. This introduces concerns regarding scalability, latency, and the robustness of inter-slice coordination. In practice, implementing such orchestration requires standardized interfaces between the slicing controller, management plane, and the data plane (many of which are still evolving within 3GPP and ETSI specifications).

The *Slice Operation Optimization Level* also presents operational challenges, particularly related to workload forecasting and dynamic scaling. Techniques based on machine learning or workload analysis require large datasets, extensive tuning, and can introduce computational overhead that partially offsets energy savings. Moreover, dynamically managing sleep modes or resource throttling demands fine-grained control at the hypervisor, container, or VNF level (capabilities that may not be uniformly supported across commercial platforms).

Lastly, the lack of unified evaluation metrics and benchmarking frameworks for energy efficiency in NS makes it

TABLE 5. NS and its prospective influence in sustainability pillars.

Sustainability Pillar	NS Contribution	Mechanism/Strategy
Environmental	Reduces overall energy consumption and emissions	• Dynamic VNF placement and scaling • Energy-aware orchestration (e.g., sleep modes) • Efficient RAN usage
	Supports integration with renewable-powered infrastructure	• Edge/cloud nodes powered by solar/wind energy • Load-aware function scheduling
Economic	Lowers OPEX through infrastructure sharing	• Multi-tenancy • NSaaS • On-demand resource allocation
	Enables new business models for vertical industries	• SLA-driven services • Custom slices for IoT, health, etc.
Social	Improves access to digital services in underserved areas	• Rural FWA using low-cost NS solutions
	Enhances quality and availability of public services	• Dedicated slices for healthcare, education, public safety

difficult for operators to assess the cost-benefit ratio of deploying such strategies. Addressing these implementation barriers will require collaboration between industry stakeholders, standardization bodies, and the research community to bridge the gap between theoretical advancements and practical deployment.

It is worth mentioning that we did not find any technique (or approach combining multiple techniques) that suits all three levels of the pyramid simultaneously. In 2016, the authors in [7] pointed out the need for a holistic approach that combined multiple energy-efficient techniques. However, little has changed since then. Our understanding is that the community is concerned about very specific challenges. In this sense, we believe that combining multiple techniques and technologies from each of the pyramid levels is an important step in understanding their final impact and overall benefits.

Developing such a holistic approach would require the creation of cross-layer optimization frameworks that dynamically adjust the placement of network functions, routing policies, and real-time resource allocation based on energy constraints. To achieve this, recent studies investigate the usage of AI-driven digital twins towards 6G networks [83], [84], [85]. In this sense, a digital twin can be understood as a virtual replica of the physical network continuously used to monitor, predict, and optimize resource allocation and energy consumption based on real-time traffic demand and environmental conditions. DRL is often integrated with digital twins to adapt slicing decisions, with the aim of optimizing resource usage and maintaining QoS constraints. However, integrating digital twins into energy-efficient network slicing remains an open challenge, particularly regarding scalability, interoperability, and real-time decision-making. Future research should focus on developing efficient AI models that can predict and optimize network energy consumption with low computational overhead.

Another observation worth highlighting is that current energy-efficient slicing approaches often assume a centralized control plane managed by a single network operator. However, real-world deployments will involve federated energy-aware resource allocation, requiring coordination between multiple infrastructure providers, such as Mobile Network Operators (MNOs), Mobile Virtual Network Operators (MVNOs), and private networks from different geographical regions. Therefore, multi-tenancy and multi-domain energy efficiency are still challenges, not only technical but also from a standardization perspective.

Finally, the energy efficiency research topic is directly related to sustainability contributions. In this regard, to the best of our knowledge, the current state-of-the-art lacks a comprehensive study on the contributions and overall impacts that Network Slicing (NS) may bring to communication networks. We recognize that the ongoing evolution of these networks can have profound implications for human society.

Thus, in this discussion, we aim to go beyond purely technical aspects of energy-saving mechanisms. We seek to emphasize the importance of long-term sustainability, highlighting that this topic remains underexplored and offers ample opportunities for impactful research. In this regard, future studies could examine not only the direct contributions of NS, but also, more broadly, the sustainability implications of advancements in communication infrastructures. To contribute to this agenda, we outline several potential contributions of NS, particularly in relation to the environmental, economic, and social pillars of sustainability. These contributions, along with the mechanisms and strategies identified in this work, are summarized in Table 5.

VII. FINAL CONSIDERATIONS

This survey considers how we can contribute to addressing the increase of energy efficiency in network slicing, with the goal of reducing ICT's carbon footprint.

The contributions presented initially include the definition of a new approach-driven taxonomy for energy-efficient network slicing. This taxonomy contributes to a more general understanding of what is being done by the specific energy-efficient solutions provided by the research community. Instead of focusing on the end algorithm or method used, the taxonomy places the proposed solutions for energy-efficient network slicing on different optimization levels that reflect the main approach used (infrastructure, topology, or slice). Concomitantly, this survey analyzes the related papers and maps them on the proposed taxonomy.

One relevant trend identified by our taxonomic classification is that energy efficiency in network slicing is gradually becoming a systemic approach. We mean that existing solutions tend to consider multiple optimization-level methods, ranging from the decision to place resources in the infrastructure to optimizing resource usage at the slice optimization level.

This suggests that a more holistic approach is the best way to save energy and reduce carbon footprint. This trend also suggests that energy-efficient systems will involve various subsystems, and machine-learning approaches will have to be integrated learning from distributed sources of knowledge. This is indeed a new challenge to be addressed by the research community.

LIST OF ACRONYMS

3GPP 3rd Generation Partnership Project
AI Artificial Intelligence
ARIMA Autoregressive Integrated Moving Average
B5G Beyond 5G
BCD block coordinate descent
BS Base Station
C-RAN Cloud-RAN
CC Cognitive Cycles
CN Core Network
CT Communication Technology
CU Central Unit
DDQN Double Deep Q-Network
DL Deep Learning
DRL Deep Reinforcement Learning
DNN Deep Neural Network
DU Distributed Unit
E2E end-to-end
EE Energy Efficiency
EESO Energy Efficient System-resource Optimization
eMBB enhanced Mobile Broadband
ETS Exponential Smoothing
FL Federated Learning
F-AP Fog Radio Access Point
F-RAN Fog-RAN
F-UE Fog-UE
FWA Fixed Wireless Access
GBR Gradient Boosting Regressor
GHG Greenhouse Gas
H-CRAN Heterogeneous Cloud Radio Access Network
HetNet Heterogeneous Network
H-NOMA heterogeneous non-orthogonal multiple access
ICT Information and Communication Technology
ILP Integer Linear Programming
IoT Internet of Things
ISP Internet Service Provider
K8s Kubernetes
KPI Key Performance Indicator
MAB Multi-Armed Bandit
MANO Management and Orchestration
MBLP Mixed Binary Linear Program
MEC Multi-access Edge Computing
MILFP mixed-integer linear fractional optimization problem
mIoT massive IoT
MILP mixed-integer linear programming
MINLP mixed-integer non-linear programming
mMTC Machine Type Communication
MNO Mobile Network Operator
MVNO Mobile Virtual Network Operator
ML Machine Learning
MVNO Mobile Virtual Network Operator
NFV Network Function Virtualization
NGMN Next-Generation Mobile Networks
NOMA Non-Orthogonal Multiple Access
NS Network Slicing
NSaaS Network Slice-as-a-Service
NSMF Network Slice Management Function
NSP Network Service Provider
OMA Orthogonal Multiple Access
O-RAN Open RAN
OPEX Operating expenses
PRB Physical Resource Block
QoE Quality of Experience
QoS Quality of Service
SFC service function chain
UE User Equipment
URLLC Ultra-Reliable and Low-Latency Communications

RAN	Radio Access Network
RB	Resource Block
RFR	Random Forest Regressor
RL	Reinforcement Learning
RMEC	Radio Multi-access Edge Computing
SACRA	Slicing Aware Clustering and Resource Allocation
SCA	Successive Convex Approximation
SDN	Software-Defined Network
SLA	Service Level Agreement
SP	Slice Provider
TL	Transfer Learning
UE	User Equipment
VNE	Virtual Network Embedding
VNF	Virtualized Network Function
VNFD	Virtualized Network Function Descriptor
VR	Virtual Reality
vRAN	Virtual RAN

REFERENCES

[1] C. Freitag, M. Berners-Lee, K. Widdicks, B. Knowles, G. S. Blair, and A. Friday, ''The real climate and transformative impact of ICT: A critique of estimates, trends, and regulations,'' *Patterns*, vol. 3, no. 8, Aug. 2022, Art. no. 100576.

[2] A. S. G. Andrae, ''New perspectives on Internet electricity use in 2030,'' *Eng. Appl. Sci. Lett.*, vol. 3, no. 2, pp. 19–31, Dec. 2020.

[3] A. Andrae and T. Edler, ''On global electricity usage of communication technology: Trends to 2030,'' *Challenges*, vol. 6, no. 1, pp. 117–157, Apr. 2015. [Online]. Available: https://www.mdpi.com/2078-1547/6/1/117

[4] J. Lorincz, A. Capone, and J. Wu, ''Greener, energy-efficient and sustainable networks: State-of-the-art and new trends,'' *Sensors*, vol. 19, no. 22, p. 4864, Nov. 2019.

[5] J. S. B. Martins et al., ''Enhancing network slicing architectures with machine learning, security, sustainability and experimental networks integration,'' *IEEE Access*, vol. 11, pp. 69144–69163, 2023.

[6] O. Alamu, A. Gbenga-Ilori, M. Adelabu, A. Imoize, and O. Ladipo, ''Energy efficiency techniques in ultra-dense wireless heterogeneous networks: An overview and outlook,'' *Eng. Sci. Technol., Int. J.*, vol. 23, no. 6, pp. 1308–1326, Dec. 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2215098619328745

[7] S. Buzzi, T. E. Klein, H. V. Poor, C. Yang, and A. Zappone, ''A survey of energy-efficient techniques for 5G networks and challenges ahead,'' *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 697–709, Apr. 2016.

[8] L. M. P. Larsen, H. L. Christiansen, S. Ruepp, and M. S. Berger, ''Toward greener 5G and beyond radio access networks—A survey,'' *IEEE Open J. Commun. Soc.*, vol. 4, pp. 768–797, 2023.

[9] I. Setiawan, B. Kar, and S.-H. Shen, ''Energy-efficient softwarized networks: A survey,'' *Trans. Netw. Service Manage.*, vol. 2023, pp. 1–21, Jan. 2023.

[10] NGMN Alliance. (Feb. 2015). *5G White Paper, Version 1.0*. [Online]. Available: https://ngmn.org/wp-content/uploads/NGMN5GWhitePaperV10.pdf

[11] M. Iwamura, ''NGMN view on 5G architecture,'' in *Proc. IEEE 81st Veh. Technol. Conf. (VTC Spring)*, May 2015, pp. 1–5.

[12] 3GPP. (2023). *Management and Orchestration; Concepts, Use Cases and Requirements (Release 18)*. Accessed: Mar. 12, 2024. [Online]. Available: https://www.3gpp.org/ftp/Specs/archive/28series/28.530/

[13] W. Lee, T. Na, and J. Kim, ''How to create a network slice?—A 5G core network perspective,'' in *Proc. 21st Int. Conf. Adv. Commun. Technol. (ICACT)*, Feb. 2019, pp. 616–619.

[14] N. F. Saraiva de Sousa, D. A. Lachos Perez, R. V. Rosa, M. A. S. Santos, and C. Esteve Rothenberg, ''Network service orchestration: A survey,'' *Comput. Commun.*, vols. 142–143, pp. 69–94, Jun. 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0140366418309502

[15] 5GPP. (Feb. 2015). *5G Vision: The Next Generation of Communication Networks and Services*. [Online]. Available: http://5g-ppp.eu/wp-content/uploads/2015/02/5G-Vision-Brochure-v1.pdf

[16] X. Foukas, G. Patounas, A. Elmokashfi, and M. K. Marina, ''Network slicing in 5G: Survey and challenges,'' *IEEE Commun. Mag.*, vol. 55, no. 5, pp. 94–100, May 2017.

[17] A. Donatti, S. L. Correa, J. S. B. Martins, A. Abelem, C. B. Both, F. Silva, J. A. Suruagy, R. Pasquini, R. Moreira, K. V. Cardoso, and T. C. Carvalho, ''Survey on machine learning-enabled network slicing: Covering the entire life cycle,'' *IEEE Trans. Netw. Service Manage.*, vol. 21, no. 3, pp. 1–18, Mar. 2023.

[18] H. P. Phyu, D. Naboulsi, and R. Stanica, ''Machine learning in network slicing—A survey,'' *IEEE Access*, vol. 11, pp. 39123–39153, 2023.

[19] X. Shen, J. Gao, W. Wu, K. Lyu, M. Li, W. Zhuang, X. Li, and J. Rao, ''AI-assisted network-slicing based next-generation wireless networks,'' *IEEE Open J. Veh. Technol.*, vol. 1, pp. 45–66, 2020.

[20] R. Moreira, F. O. Silva, T. C. M. B. Carvalho, and J. S. B. Martins, ''Intelligent data-driven architectural features orchestration for network slicing,'' in *Proc. Int. Workshop Adv. ICT Infrastruct. Services*, Jan. 2024, pp. 1–12.

[21] H. Chergui, L. Blanco, L. A. Garrido, K. Ramantas, S. Kuklinski, A. Ksentini, and C. Verikoukis, ''Zero-touch AI-driven distributed management for energy-efficient 6G massive network slicing,'' *IEEE Netw.*, vol. 35, no. 6, pp. 43–49, Nov. 2021.

[22] H. Khan, M. M. Butt, S. Samarakoon, P. Sehier, and M. Bennis, ''Deep learning assisted CSI estimation for joint URLLC and eMBB resource allocation,'' in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, Jun. 2020, pp. 1–6.

[23] J. Tang, Y. Duan, Y. Zhou, and J. Jin, ''Distributed slice selection-based computation offloading for intelligent vehicular networks,'' *IEEE Open J. Veh. Technol.*, vol. 2, pp. 261–271, 2021.

[24] J. A. Ayala-Romero, A. Garcia-Saavedra, M. Gramaglia, X. Costa-Pérez, A. Banchs, and J. J. Alcaraz, ''VrAIn: Deep learning based orchestration for computing and radio resources in vRANs,'' *IEEE Trans. Mobile Comput.*, vol. 21, no. 7, pp. 2652–2670, Jul. 2022.

[25] N. Yu, Z. Song, H. Du, H. Huang, and X. Jia, ''Multi-resource allocation in cloud radio access networks,'' in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2017, pp. 1–6.

[26] R. Moreira, J. S. B. Martins, T. C. M. B. Carvalho, and F. D. O. Silva, ''On enhancing network slicing life-cycle through an AI-native orchestration architecture,'' in *Advanced Information Networking and Applications* (Lecture Notes in Networks and Systems), L. Barolli, Ed., Cham, Switzerland: Springer, 2023, pp. 124–136.

[27] E. T. Martínez Beltrán, M. Q. Pérez, P. M. S. Sánchez, S. L. Bernal, G. Bovet, M. G. Pérez, G. M. Pérez, and A. H. Celdrán, ''Decentralized federated learning: Fundamentals, state of the art, frameworks, trends, and challenges,'' *IEEE Commun. Surveys Tuts.*, vol. 25, no. 4, pp. 2983–3013, 4th Quart., 2023.

[28] Y. Azimi, S. Yousefi, H. Kalbkhani, and T. Kunz, ''Energy-efficient deep reinforcement learning assisted resource allocation for 5G-RAN slicing,'' *IEEE Trans. Veh. Technol.*, vol. 71, no. 1, pp. 856–871, Jan. 2022.

[29] L. Blanco, E. Zeydan, S. Barrachina-Muñoz, F. Rezazadeh, L. Vettori, and J. Mangues-Bafalluy, ''A novel approach for scalable and sustainable 6G networks,'' *IEEE Open J. Commun. Soc.*, vol. 5, pp. 1673–1692, 2024.

[30] H. Chergui, L. Blanco, and C. Verikoukis, ''Statistical federated learning for beyond 5G SLA-constrained RAN slicing,'' *IEEE Trans. Wireless Commun.*, vol. 21, no. 3, pp. 2066–2076, Mar. 2022.

[31] A. Thantharate, A. V. Tondwalkar, C. Beard, and A. Kwasinski, ''ECO6G: Energy and cost analysis for network slicing deployment in beyond 5G networks,'' *Sensors*, vol. 22, no. 22, p. 8614, Nov. 2022. [Online]. Available: https://www.mdpi.com/1424-8220/22/22/8614

[32] R. Bolla, R. Bruschi, F. Davoli, L. Ivaldi, C. Lombardo, and B. Siccardi, ''An AI framework for fostering 6G towards energy efficiency,'' in *Proc. 61st FITCE Int. Congr. Future Telecommun., Infrastruct. Sustainability (FITCE)*, Sep. 2022, pp. 1–6.

[33] S. Sundaramoorthy, D. Kamath, S. Nimbalkar, C. Price, T. Wenning, and J. Cresko, ''Energy efficiency as a foundational technology pillar for industrial decarbonization,'' *Sustainability*, vol. 15, no. 12, p. 9487, Jun. 2023.

[34] F. S. Hafez, B. Sa'di, M. Safa-Gamal, Y. H. Taufiq-Yap, M. Alrifaey, M. Seyedmahmoudian, A. Stojcevski, B. Horan, and S. Mekhilef, ''Energy efficiency in sustainable buildings: A systematic review with taxonomy, challenges, motivations, methodological aspects, recommendations, and pathways for future research,'' *Energy Strategy Rev.*, vol. 45, Jan. 2023, Art. no. 101013.

[35] M. G. Patterson, "What is energy efficiency? Concepts, indicators and methodological issues," *Energy Policy*, vol. 24, no. 5, pp. 377–390, May 1996.

[36] Y. Anser, J.-L. Grimault, S. Bouzefrane, and C. Gaber, "Energy-aware service level agreements in 5G NFV architecture," in *Proc. 8th Int. Conf. Future Internet Things Cloud (FiCloud)*, Aug. 2021, pp. 377–382.

[37] R. W. Wimbadi and R. Djalante, "From decarbonization to low carbon development and transition: A systematic literature review of the conceptualization of moving toward net-zero carbon dioxide emission (1995–2019)," *J. Cleaner Prod.*, vol. 256, May 2020, Art. no. 120307.

[38] UNDESA. *Sustainable Development Goals: The 17 UN Goals for 2030*. [Online]. Available: https://sdgs.un.org/goals

[39] A. T. Rosário and J. C. Dias, "The new digital economy and sustainability: Challenges and opportunities," *Sustainability*, vol. 15, no. 14, p. 10902, Jul. 2023.

[40] Z. Wang, H.-T. Liao, J. Lou, and Y. Liu, "Making cyberspace towards sustainability a scientometric review for a cyberspace that enables green and digital transformation," in *Proc. Int. Conf. Cyberspace Innov. Adv. Technol.*, New York, NY, USA, Jan. 2021, pp. 394–400.

[41] L. Farhan, R. S. Hameed, A. S. Ahmed, A. H. Fadel, W. Gheth, L. Alzubaidi, M. A. Fadhel, and M. Al-Amidie, "Energy efficiency for green Internet of Things (IoT) networks: A survey," *Network*, vol. 1, no. 3, pp. 279–314, Nov. 2021.

[42] J.-L. Hu, Y.-C. Chen, and Y.-P. Yang, "The development and issues of energy-ICT: A review of literature with economic and managerial viewpoints," *Energies*, vol. 15, no. 2, p. 594, Jan. 2022.

[43] Y. Jin, Y. Wen, and Q. Chen, "Energy efficiency and server virtualization in data centers: An empirical investigation," in *Proc. IEEE INFOCOM Workshops*, Mar. 2012, pp. 133–138.

[44] H. J. Loschi, J. Leon, Y. Iano, E. R. Filho, F. D. Conte, T. C. Lustosa, and P. O. Freitas, "Energy efficiency in smart grid: A prospective study on energy management systems," *Smart Grid Renew. Energy*, vol. 6, no. 8, pp. 250–259, 2015.

[45] X. Li, M. Samaka, H. A. Chan, D. Bhamare, L. Gupta, C. Guo, and R. Jain, "Network slicing for 5G: Challenges and opportunities," *IEEE Internet Comput.*, vol. 21, no. 5, pp. 20–27, May 2017.

[46] C. Najjuuko, G. K. Ayebare, R. Lukanga, E. Mugume, and D. Okello, "A survey of 5G for rural broadband connectivity," in *Proc. IST-Africa Conf. (IST-Africa)*, May 2021, pp. 1–10.

[47] A. Markhasin, "Fundamentals of the extremely green, flexible, and profitable 5G M2M ubiquitous communications for remote e-healthcare and other social e-applications," in *Proc. Int. Multi-Conf. Eng., Comput. Inf. Sci. (SIBIRCON)*, Sep. 2017, pp. 292–297.

[48] Y. Lin, L. Shao, Z. Zhu, Q. Wang, and R. K. Sabhikhi, "Wireless network cloud: Architecture and system requirements," *IBM J. Res. Develop.*, vol. 54, no. 1, pp. 4:1–4:12, Jan. 2010.

[49] J. U. Patel, S. J. Guercio, A. E. Bruno, M. D. Jones, and T. R. Furlani, "Implementing green technologies and practices in a high performance computing center," in *Proc. Int. Green Comput. Conf.*, Jun. 2013, pp. 1–8.

[50] M. M. Vishwakarma and M. Manoria, "A robust algorithm for capacity management to improve response time in virtualized data centers," in *Proc. Int. Conf. Comput. Intell. Commun. Netw. (CICN)*, Dec. 2015, pp. 731–735.

[51] J. L. Bolonhezi Dias, L. B. De Almeida, and L. C. Pessoa Albini, "Reducing Hadoop 3.X energy consumption through energy efficient Ethernet," in *Proc. IEEE Int. Conf. Big Data Smart Comput. (BigComp)*, Jan. 2022, pp. 9–14.

[52] H. Xiang, M. Peng, Y. Sun, and S. Yan, "Mode selection and resource allocation in sliced fog radio access networks: A reinforcement learning approach," *IEEE Trans. Veh. Technol.*, vol. 69, no. 4, pp. 4271–4284, Apr. 2020.

[53] H. P. Phyu, D. Naboulsi, R. Stanica, and G. Poitau, "Towards energy efficiency in RAN network slicing," in *Proc. IEEE 48th Conf. Local Comput. Netw. (LCN)*, Oct. 2023, pp. 1–9.

[54] N. Sen, "Towards energy efficient functional split and baseband function placement for 5G RAN," in *Proc. IEEE 9th Int. Conf. Netw. Softwarization (NetSoft)*, Jun. 2023, pp. 237–241.

[55] H.-W. Kao and E. H. Wu, "QoE sustainability on 5G and beyond 5G networks," *IEEE Wireless Commun.*, vol. 30, no. 1, pp. 118–125, Feb. 2023.

[56] M. A. Hossain and N. Ansari, "Network slicing for NOMA-enabled edge computing," *IEEE Trans. Cloud Comput.*, vol. 11, no. 1, pp. 811–821, Jan. 2023.

[57] A. Hossain and N. Ansari, "5G multi-band numerology-based TDD RAN slicing for throughput and latency sensitive services," *IEEE Trans. Mobile Comput.*, vol. 22, no. 3, pp. 1263–1274, Mar. 2023.

[58] S. Ismail, F. D'Andreagiovanni, H. Lakhlef, and Y. Imine, "Recent advances on 5G resource allocation problem using PD-NOMA," in *Proc. Int. Symp. Netw., Comput. Commun. (ISNCC)*, Oct. 2020, pp. 1–7.

[59] B. Liu, P. Zhu, J. Li, D. Wang, and X. You, "Energy-efficient optimization in distributed massive MIMO systems for slicing eMBB and URLLC services," *IEEE Trans. Veh. Technol.*, vol. 72, no. 8, pp. 10473–10487, Aug. 2023.

[60] S. Zou, H. Yu, W. Wang, and W. Ni, "An energy-efficient framework in radio multi-access edge cloud by automatically generating network slice," *IEEE Internet Things Mag.*, vol. 5, no. 3, pp. 128–133, Sep. 2022.

[61] A. Ndikumana, K. K. Nguyen, and M. Cheriet, "Renewable energy powered and open RAN-based architecture for 5G fixed wireless access provisioning in rural areas," *IEEE Trans. Green Commun. Netw.*, vol. 8, no. 3, pp. 994–1007, Sep. 2024.

[62] W.-K. Chen, Y.-F. Liu, A. De Domenico, and Z.-Q. Luo, "Network slicing for service-oriented networks with flexible routing and guaranteed E2E latency," in *Proc. IEEE 21st Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, May 2020, pp. 1–5.

[63] S. Mishra and C. Hota, "Joint QoS and energy-efficient resource allocation and scheduling in 5G network slicing," *Comput. Commun.*, vol. 202, pp. 110–123, Mar. 2023.

[64] A. Mesodiakaki, M. Gatzianas, G. Kalfas, F. Moscatelli, G. Landi, N. Ciulli, and L. Lossi, "5G-COMPLETE: End-to-end resource allocation in highly heterogeneous beyond 5G networks," in *Proc. IEEE 4th 5G World Forum (5GWF)*, Oct. 2021, pp. 412–417.

[65] J. Ma, C. Pan, C. Yin, and X. Li, "Slice-aware resource management in SDN enabled heterogeneous cellular networks," in *Proc. IEEE/CIC Int. Conf. Commun. China (ICCC)*, Aug. 2019, pp. 869–874.

[66] Y. Xu, Z. He, Y. Zhang, and W. Zhou, "Energy-efficient resource allocation for slicing-enabled multi-access edge computing," in *Proc. IEEE Smartworld, Ubiquitous Intell. Comput., Scalable Comput. Commun., Digit. Twin, Privacy Comput., Metaverse, Auto. Trusted Vehicles (SmartWorld/UIC/ScalCom/DigitalTwin/PriComp/Meta)*, Dec. 2022, pp. 698–705.

[67] P. Han, Y. Liu, X. Zhang, and L. Guo, "Energy-efficient service placement based on equivalent bandwidth in cell zooming enabled mobile edge cloud networks," *IEEE Trans. Veh. Technol.*, vol. 71, no. 11, pp. 12275–12290, Nov. 2022.

[68] S. O. Oladejo and O. Falowo, "An energy-efficient resource allocation scheme for 5G slice networks," in *Proc. Southern Afr. Telecommun. Netw. Appl. Conf.*, 2019, pp. 1–6.

[69] Q. Chen, X. Xu, and H. Jiang, "Online green communication scheduling for sliced unlicensed heterogeneous networks," *IEEE Trans. Veh. Technol.*, vol. 70, no. 10, pp. 10657–10670, Oct. 2021.

[70] D. Basu, A. Jain, U. Ghosh, and R. Datta, "QoS-aware dynamic network slicing and VNF embedding in softwarized 5G networks," in *Proc. Nat. Conf. Commun. (NCC)*, May 2022, pp. 100–105.

[71] H. Cao, J. Du, H. Zhao, D. X. Luo, N. Kumar, L. Yang, and F. R. Yu, "Toward tailored resource allocation of slices in 6G networks with softwarization and virtualization," *IEEE Internet Things J.*, vol. 9, no. 9, pp. 6623–6637, May 2022.

[72] J. Zhou, W. Zhao, and S. Chen, "Dynamic network slice scaling assisted by prediction in 5G network," *IEEE Access*, vol. 8, pp. 133700–133712, 2020.

[73] H.-M. Chen, S.-Y. Chen, S.-K. Wang, and J.-Q. Chen, "Designing a reinforcement learning approach for the NFV orchestration system with energy saving optimization," in *Proc. 8th Int. Conf. Appl. Syst. Innov. (ICASI)*, Apr. 2022, pp. 98–101.

[74] M. A. Aba, M. Kassis, M. Elkael, A. Araldo, A. A. Khansa, H. Castel-Taleb, and B. Jouaber, "Efficient network slicing orchestrator for 5G networks using a genetic algorithm-based scheduler with Kubernetes: Experimental insights," in *Proc. IEEE 10th Int. Conf. Netw. Softwarization (NetSoft)*, Jun. 2024, pp. 82–90.

[75] C. Xing, Y. L. Lee, and Y. C. Chang, "Joint baseband and radio resource allocation for 5G network slicing in H-CRANs," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, May 2022, pp. 891–896.

[76] J. Miao, X. Chai, X. Song, and T. Song, "A DDQN-based energy-efficient resource allocation scheme for low-latency V2V communication," in *Proc. IEEE 5th Int. Elect. Energy Conf. (CIEEC)*, May 2022, pp. 53–58.

[77] S. S. Jayanthi, Y. L. Lee, and Y. C. Chang, "User association for multi-tenant heterogeneous network slicing using genetic algorithm," in *Proc. 8th Int. Conf. Comput. Commun. Eng. (ICCCE)*, Jun. 2021, pp. 326–330.

[78] Y. K. Tun, D. H. Kim, M. Alsenwi, N. H. Tran, Z. Han, and C. S. Hong, "Energy efficient communication and computation resource slicing for eMBB and URLLC coexistence in 5G and beyond," *IEEE Access*, vol. 8, pp. 136024–136035, 2020.

[79] M. Laroui, M. A. Cherif, H. I. Khedher, H. Moungla, and H. Afifi, "Scalable and cost efficient resource allocation algorithms using deep reinforcement learning," in *Proc. Int. Wireless Commun. Mobile Comput. (IWCMC)*, Jun. 2020, pp. 946–951.

[80] S. Ravindran, S. Chaudhuri, J. Bapat, and D. Das, "EESO: Energy efficient system-resource optimization of multi-sub-slice-connected user in 5G RAN," in *Proc. IEEE Int. Conf. Electron., Comput. Commun. Technol. (CONECCT)*, Jul. 2020, pp. 1–6.

[81] Y. Zhang, L. Zhao, K. Liang, G. Zheng, and K.-C. Chen, "Energy efficiency and delay optimization of virtual slicing of fog radio access network," *IEEE Internet Things J.*, vol. 10, no. 3, pp. 2297–2313, Feb. 2023.

[82] D.-D. Tran, V. N. Ha, S. K. Sharma, T. T. Nguyen, S. Chatzinotas, and P. Popovski, "Energy-efficient NOMA for 5G heterogeneous services: A joint optimization and deep reinforcement learning approach," *IEEE Trans. Commun.*, vol. 73, no. 4, pp. 2448–2465, Apr. 2025.

[83] Z. Zhang, Y. Huang, C. Zhang, Q. Zheng, L. Yang, and X. You, "Digital twin-enhanced deep reinforcement learning for resource management in networks slicing," *IEEE Trans. Commun.*, vol. 72, no. 10, pp. 6209–6224, Oct. 2024.

[84] M. Yaqoob, R. Trestian, M. Tatipamula, and H. X. Nguyen, "Digital-twin-driven end-to-end network slicing toward 6G," *IEEE Internet Comput.*, vol. 28, no. 2, pp. 47–55, Mar. 2024, doi: 10.1109/MIC.2023.3332252.

[85] J. Wang, J. Li, and J. Liu, "Digital twin-assisted flexible slice admission control for 5G core network: A deep reinforcement learning approach," *Future Gener. Comput. Syst.*, vol. 153, pp. 467–476, Apr. 2024, doi: 10.1016/j.future.2023.12.018.

ADNEI WILLIAN DONATTI received the B.Tech. degree in information technology and the M.Sc. degree in applied computing from Santa Catarina State University (UDESC). He is currently pursuing the Ph.D. degree in computer engineering with the Escola Politécnica, Universidade de São Paulo (USP), under the Graduate Program in Electrical Engineering (PPGEE). He conducts research on energy-efficient network slicing with a focus on intelligent virtual network function (VNF) placement and orchestration, leveraging deep reinforcement learning (DRL) techniques. His recent work explores scalable, cross-layer optimization frameworks for 5G and beyond, integrating AI-based decision systems and sustainability-aware policies. He is actively involved in the Smart Future Internet Infrastructure (SFI^2) project and leads the development of reproducible pipelines for large-scale genotype and phenotype data analysis using Spark and R on cloud platforms. His research interests include network softwarization, cloud computing, distributed systems, and sustainable computing.

MARCIA CRISTINA MACHADO received the degree in computer science from the University of Guarulhos, in 1994, the master's degree in information systems from the Pontifical Catholic University of São Paulo, in 2002, the master's degree in business administration from the University Nove de Julho, in 2016, and the Ph.D. degree in computer engineering and digital systems from the Universidade de São Paulo. She is currently a Postdoctoral Researcher at the Polytechnic School, Universidade de São Paulo, researching energy efficiency, sustainability indicators, and machine learning applications. She works as a Guest Professor on M.B.A. courses at POLI-USP and FIA-FEA/USP.

MARVIN ALEXANDER LOPEZ MARTINEZ received the bachelor's degree in international business from the Technological University of El Salvador (UTEC) and the M.Sc. degree in information technologies from Nikola Vaptsarov Naval Academy (NVNA). He is currently pursuing the Ph.D. degree with the Universidade Federal de Pernambuco (UFPE), working on the power consumption of network slicing in the SFI2 project. His research interests include slice energy consumption, slice energy efficiency, and energy consumption modeling.

SABINO ROGÉRIO S. ANTUNES received the degree from the Universidade Salgado de Oliveira (UNIVERSO), in 2006, with a specialization in network systems security from Faculdade Santa Maria (FSM), in 2011, and the M.Sc. degree in computer science from the Universidade Federal de Pernambuco (UFPE), Brazil, in 2018, where he is currently pursuing the Ph.D. degree. He worked as a Lecturer in computer networks at Universidade Joaquim Nabuco, from 2011 to 2019. He currently holds the position of IT technician at UFPE. His research interests include cloud computing, computer networks, mobile networks, and network slicing.

ELI CARLOS FIGUEIREDO SOUZA, photograph and biography not available at the time of publication.

SAND L. CORRÊA received the bachelor's degree in computer science from the Universidade Federal de Goiás (UFG), in 1994, the M.Sc. degree in computer science from the State University of Campinas (Unicamp), and the D.Sc. degree in informatics from the Pontifical Catholic University of Rio de Janeiro (PUC-Rio), in 2011. Since 2010, she has been an Associate Professor at the Institute of Informatics, UFG.

TIAGO C. FERRETO (Member, IEEE) received the Ph.D. degree in computer science from the Computer Science Department, Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS), Brazil, in 2010, with a Sandwich Ph.D. Internship at Technische Universität Berlin, Germany, from 2007 to 2008. He is currently an Associate Professor of computer science at PUCRS, where he is also the Head of the Research Group on Networking, Infrastructure, and Cloud Computing (GRIN) and the High-Performance Computing Laboratory (LAD). His research interests include cloud and edge computing, IT infrastructure management, computer networks, high-performance computing, and big data.

JOSÉ AUGUSTO SURUAGY (Life Senior Member, IEEE) received the Ph.D. degree in computer science from the University of California at Los Angeles (UCLA), in 1990. He is currently a Professor and a Researcher at the CESAR School. Has experience in computer science, focusing on telecomputing, acting on the following subjects: computer networks, measurement, quality of service, performance evaluation, and cybersecurity. In 2014, he received the Computer Network Brazilian Symposium Distinguished Award for his scientific contributions in the areas of computer networks and distributed systems, for his participation in the symposium activities, and for his services that benefited the community. In 2019, he received the Newton Faller Award from Brazilian Computer Society, in recognition of his brilliant collaboration and dedication to the exercise of multiple activities of great interest to Brazilian Computing Community, in particular, his work with the Forum of Computer Science Graduate Chairs.

JOBERTO S. B. MARTINS (Life Senior Member, IEEE) received the Ph.D. degree in computer science from Université Pierre et Marie Curie (UPMC), Paris, in 1986. He was a Postdoctoral Researcher with ICSI—Berkeley University, in 1995, and a Postdoctoral Senior Researcher at Paris Saclay University, France, in 2016. He has been an International Professor at Hochschule für Technik und Wirtschaft des Saarlandes (HTW), Germany, since 2004. He is currently a Full Professor in computer science with Universidade Salvador (UNIFACS) and the Director of the NUPERC Research Group. His research interests include network slicing, resource orchestration, machine learning, and smart city. He is a key speaker, a teacher, and an invited lecturer at various international congresses and companies in Brazil and Europe.

TEREZA CRISTINA CARVALHO is currently an Associate Professor with Escola Politécnica, Universidade de São Paulo (USP), and a Visiting Professor with Université Paris 1 Panthéon-Sorbonne. She has been the Founder and a general Coordinator of the Laboratory of Sustainability on ITC (LASSU), since 2010, and the Center for Reuse and Discard of Informatics Residuals (CEDIR), USP, since 2009. She was a former Assessor with Information Technology Coordination (CTI), USP, from 2010 to 2013, and the Director of Electronic Computing Center (CCE), USP, from 2006 to 2010. She was a Sloan Fellow with Massachusetts Institute of Technology (MIT), in 2002. She has coordinated international and national research and development projects, since 2000, in green computing, cloud computing, IT energy efficiency, IT governance, digital technologies applied to the Amazon Production Chains, Waste Electrical and Electronic Equipment (WEEE), future internet, scientific DMZ, and security. She holds several international patents.