Energy Efficient Technologies for Virtualized Cloud Data Center: A Systematic Mapping Study (SMS) And Systematic Literature Review (SLR)

Sarwan Soomro[1], Ishan Sharma[2], Sheeraz Gul[3]

[1]University of Sindh, Jamshoro, Pakistan
[2]MSc South China University of Technology, Guangzhou, China
[3]PhD Student Northwestern Polytechnical University, Xian, China

---***---

Abstract - *Cloud computing has transformed into a daily life activity for people worldwide. It helps connect billions of users digitally. Because of massive data centers, the ever-growing energy consumption affects the environment and increases the cost for the service providers. Which then leads to a poor Service Level Agreement (SLA). This paper conducted a systematic mapping study on 74 peer-reviewed articles to evaluate energy-efficient technologies that optimize power consumption in virtualized data centers. Moreover, we proposed a characterization framework to select only relevant data before classifying it with our conceptualization map. We also distributed the studies according to the characterization criteria: a) generic attributes, b) contribution type and evaluation method, c) technological attributes, and d) quality management. The results showed that virtualization, consolidation, and workload scheduling are widely implemented techniques. In addition, results suggested that the contribution type in around 60% of the studies is based on solution and validation. The methods either involved specific experiments or comprised theoretical examples of model development. Likewise, DVFS-enabled scheduling and dynamic server consolidation methods seemed vital in saving energy in the virtualized cloud data center. In brief, we surveyed an existential need for a standardized and centralized benchmarking system for researchers to bridge the gap between industry and academia. Also, the results can help understand the current trends in this research domain, and this paper can be used as a benchmark to assess and evaluate current research progress.*

Key Words: ***Data Centers; Energy-Efficiency; Cloud Computing; Virtualization; Consolidation; Scheduling; DVFS; Evaluation; Metrics***

1. INTRODUCTION

The growth of cloud data centers has also increased the development cost, which amounts to approximately $20 billion yearly.[2]. Similarly, power consumed by data centers has also become a major concern for the service providers and the environment because of the carbon emissions caused, which have also risen exponentially in the recent past. The major contributor to the energy consumption in a data center is the IT infrastructure, which comprises servers, cooling systems, lighting, and other IT equipment. Likewise, the enterprise servers' power consumption dominates the data center's power, representing up to 60% of the overall data center consumption." [3]. IT systems in data centers play an integral role in productivity and reliability outcomes since a traditional data center comprises power-hungry servers, which can be dynamic or static power contributions based on how they function.

For example, a traditional data center consumes more power due to outdated infrastructure, poor energy management, and underlying system incompatibility issues. On the contrary, power-efficient data centers release fewer carbon emissions. In addition, they can also help cloud service providers with productivity and cost-effectiveness. Provided the proper implementation of energy-efficient techniques and quality management of physical resources is addressed. The most common methods to deal with energy consumption in a typical data center are virtual machine placement and server consolidation, especially in a virtualized data center. Modern interconnectivity is a kind of digital globalization that has attracted billions of active users worldwide. Hence, the quantity of servers in a typical data center has also increased. A typical data center comprises various components, such as a cooling system, servers, storage, etc. Moreover, servers consume more energy, leading to costs that are a big concern for Cloud Service Providers (CSPs). Throughout the 2011-2035 timeframe, forecasts suggest that the energy production for cloud services to expand by upwards of 66%.[4]. Therefore, cloud services ought to enforce energy-efficient data center management to cope with the growing energy needs for cloud services and satisfy the demands for cloud applications while ensuring the SLA and usability goals of the end-users. Besides, it is also to sustain low costs, provided that the interest in saving energy must be a concern for Cloud Service Providers, especially big corporations. [5] Given the massive energy consumption aspects of data centers, one of the major issues is that power measures are essential for ensuring energy and cost. It means determining efficient resource allocation to each client's requests to satisfy client needs and meet the service provider's business model. Most of these priorities fall into the energy management of data centers or cost reductions for the data centers because these are the cloud computing backbones. Hence, energy-saving technologies can help create a sustainable, reliable, and eco-friendly infrastructure model across various platforms to be followed as a benchmark, which would benefit CSPs and the environmental cause and end-users in the long run. Therefore, this paper successfully implemented a Systematic Mapping Study (SMS) and Systematic Literature review (S.L.R.) with a characterization framework on the broadness of energy-efficient technologies, details of energy-saving methods, and the future of such developments for cloud data centers. In addition, this study can help reflect on environmental problems caused by a growing number of electrical, power-hungry systems that suck up energy inefficiently. Our study analyses research venues and methods using a set of pre-defined, easy-to-track steps. We proposed a characterization framework that extracts invaluable information from 74 selected publications based on their relevance, soundness, and solutions. We categorized each section using a mapping criterion. In addition, solutions, examples, experiences, and evaluation methods are systematically mapped in section 4. Given the maturity of power management and energy-efficient strategies, this paper aims to classify and synthesize a comparative overview of recent research works and map an assessment of ongoing work in the data centers' energy-saving technologies. This paper also identifies the global trend, assesses the industry standards, and evaluates the emerging methodologies taxonomically in domain-specific and generic manners.

Furthermore, this undertaking has resulted in a knowledge base that is the essence of modern approaches, quality features, techniques, and best practices for Cloud Energy Efficiency, power management, and data center performance. We have systematically divided the sections for higher readability and effective analysis.

1. Firstly, our work discusses and overviews the energy consumption problem globally. Then, it maps out the energy-saving methods and power management strategies through comparison. It also highlights best practices and standard benchmarks to specify the appropriate solution and benchmarks for scientists to study.

2. We overview traditional literature on energy-saving techniques in virtualized cloud data centers. We focus on major techniques while reviewing the latest algorithms, their service quality, and real-world parameters.

3. Section 3 addresses our proposed systematic mapping study with methodological steps. Moreover, it also proposed a characterization framework, our novel approach to classify this research into various components, such as technologies, quality parameters, research methods, and other variables.

4. Section 4 presents the results comprehensively. We use a modular approach to classify research venues, methods, and significant empirical data. That includes tables, graphs, and a critical analysis of each finding. As it narrows down each technique, it highlights significant findings. And it also discusses the progress made in the area of research. Also, it reviews the promising technologies that scientists can use in the future.

5. Section 5 examines the open challenges and critical analysis of the literature and concludes the study with a comprehensive discussion.

2. Related Work

In [21], authors surveyed computational clouds using consolidation, resource allocation, and virtualization techniques. Their approach is traditional, as it needs to use more data to set a benchmark. In [22], authors reviewed 68 studies and concluded that the server infrastructure is the main energy consumer. Besides, a work by [23] conducted a systematic mapping study (SMS) on five databases, used 58 publications for their review, and found that almost all the studies were solution-based or experimental. Nevertheless, their conclusion lacked crucial parameters such as the level of deployment, execution, quality of service, and infrastructure metrics. In addition, a study undertaken by [24] concludes that using hardware frequency scaling capabilities in clusters can save energy for parallel programs during their communication times in high-performance settings. Chase et al. explain methods to reduce power usage in data centers by activating and disabling servers on demand for server systems [6]. Support for energy-efficient disc arrays has been extended [39] to identify the data center's requirement for disc power control. The work [25] proposed incorporating temperature awareness into workload distribution in data centers, coupled with emulation environments, to investigate the thermal consequences of power management [15]. Researchers have also studied the use, in conjunction with the DVFS [8], of such cluster-level reconfiguration or spare servers [32] to increase the thermal and energy efficiency of business systems. Other techniques have enforced power budgets by non-uniform allocation of power between nodes [10], blade enclosure granularity [33], or even virtual machines employing energy accounting capabilities [35]. The experimental assessments conducted in this work show that global management policies such as these are easily implemented with the right VPM rules based on the rich range of power control methods used by Virtual Power. In data center environments, the ability to handle hardware heterogeneity is an important attribute of online power management. This is apparent from previous work on heterogeneous multi-core architectures management extensions [21] and cluster environments, the latter proposing a scheduled approach to the power control of processors with varying fixed frequencies and voltages [13]. Smart strategies for distributing requests in corporate systems have used hardware heterogeneity to reduce power use [16]. There have also been investigations on using heterogeneity in hardware and the underlying power management capabilities for data centers [26]. The experiments here highlight heterogeneity support's significance in Virtual Power and further demonstrate the utility of heterogeneity awareness in power management approaches.

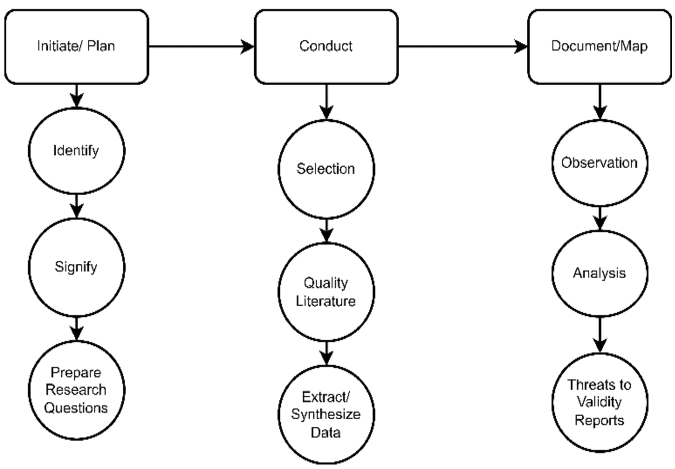
Table -1: Comparison between our Systematic Mapping Study and Others'

Ref.	Study Type	Total Studies	Search Strategy	Search Eval	Time Period	Framework	Contribution & Scope
[1]	Survey	N/A	Manual	X	N/A	X	Energy-saving technologies for two components of data centers: IT equipment and cooling systems-Mathematical support and discussion on AI based techniques-
[2]	Survey	N/A	Manual	X	N/A	X	Taxonomies for categorizing and reviewing current research, as well as an analysis and discussion of data center energy-saving technologies-
[3]	SMS	58	Snowballing	X	2017-2019	X	To render cloud data centres energy-efficient, software strategies, affected stakeholders, performance features, databases, and tools were used.
[4]	Survey	N/A	Manual Snowballing	X	N/A	X	Server consolidation strategies are divided into four categories: static, dynamic, prediction-based dynamic, and hybrid.
[5]	Survey	X	Manual	X	N/A	X	A method for evaluating the energy performance of the most critical data center domains, such as server and network equipment, using a systematic approach.
	This Work Proses:	74	Manual, Snowballing, and Database		2005-2021	Characterization	Data Center Energy Saving Technologies in Cloud: A systematic mapping study

3. RESEARCH METHODLOGY

Our systematic mapping study (SMS) follows a sequential process to categorize primary search results for an effective literature review. It comprises a set of methods tracked through systematic steps for the classification and analysis of the selected data. Our inspiration is drawn from the work presented in [24, 25], which is a three-step process for quality review; it comprises the study selection, assessment, and conclusion strategy:

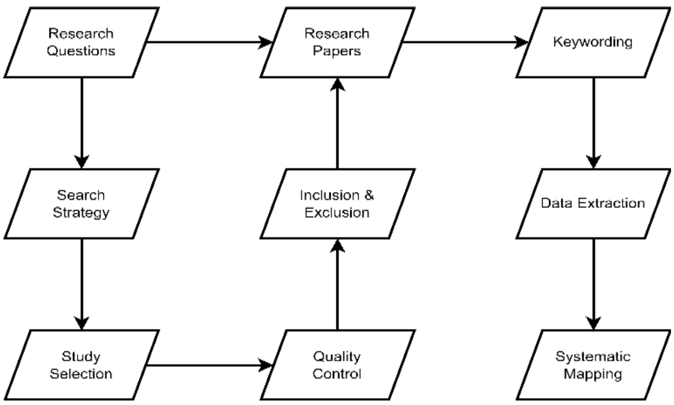
Fig -1: Three Steps Process to conduct our (SMS)



For better visualization of individual steps, we have formulated and revised some of the steps to suit our work, shown in (Fig - 2). The review follows a systematic mapping process with a predicted outcome for each action. Besides, we provide an additional categorization scheme (Fig - 4), a prerequisite to the (Table - 7), which is our conceptualization to comprehensively classify selected data. We aim to review energy-efficient technologies in a broad context and data centers' energy-saving mechanisms with a taxonomical review strategy in specific. Moreover, we review and philosophize current research into perspective by using a systematic approach. This work has been conducted on the selected studies listed in (Table - 7) with clarity of context, relevance, authenticity, and significance.

Moreover, we go through research questions and conduct a primary search on model papers based on their credibility and publication standard. Thereafter, the screening of papers is defined and keywording based on a specific search strategy using the title and abstract. Then we extract and map all the data, which is presented graphically in (Section 4). As we focus on systematic mapping study, each step we take has an outcome based on the mapping. First, we must systematically classify each step to initiate mapping to exploit our literature extraction. It is worth noting that each step this study takes has an outcome.

Fig -2: Systematic Mapping Study Process for our work



The methodological approaches demonstrate the current procedures that must be resolved or the usefulness of gathering data and answering questions. Therefore, the research question will help the reader assess the research's scope.

3.1 Research Questions

Q.1 What is the frequency of research conducted to optimize energy in cloud data centers in a broad view?

- To understand the broad activity in the research area, to classify domains, and to analyze trends-

Q.2 What is the frequency of a particular research technique as a subcategory?

- To understand applications into an individual method based on its activity and frequency. To evaluate and reflect on active journals, publishers, authors, countries, and institutions actively contributing to this research.

Q.3 What are the evaluation methods and their contributions and what is the relevance of employing such a technique and its practical implications?

- To accumulate a solid perspective into the solutions, proposals, evaluations, and validations of an individual research item and to differentiate among empirical research- To reflect on technology stacks, categorize, evaluate, and understand each technique.

Q.4 What level of quality is the study on management, service level agreements, and parameters determining the quality of service, as well as the classification of models, frameworks, and learning curves in all of those?

- To understand the quality of service, management strategies, and architectural concerns.

Q.5 What are some of the reservations, and what is the future of energy conservation in the data center?

- The central idea is to reflect on the gaps while assessing the best practices and finally come up with fruitful suggestions for the active participants in the domain-

Our search strategy is based on PICO (Population, intervention, comparison, and outcomes) suggested [24], which was developed to identify keywords and formulate search strings from research questions. We use PICO search guidelines as our primary search method due to its reliability and broad scientific citation. We apply our context for scoping and mapping energy-saving technologies in the Cloud, guided by individual steps. It is essential to have a search strategy to minimize the researcher's bias. Because a credible visualization of extracted data is only possible if the data collected is precise, brief, sound, and relevant. The quest for primary studies is usually carried out by scouring significant databases. It can also be accomplished manually by looking through conference and journal materials. We have assigned attributes derived from RQ1 to RQ5 to conceptualize our PICO search strategy, and each point is aligned with the PICO search term, which can be seen in (Table - 2).

Table -2: PICO Method

PICO Method [6]	Summary
Population	Need for the Study/Frequencies of the research venues. Frequencies of individual research items. Historical context, significance, worldview, technical, commercial, and environmental aspects. Quality of service, execution, architectural, technological, and maturity stage. Future work, conclusions, philosophies, and critical analysis.
Intervention	Categorization, extraction, validation, and analysis
Comparison	Mapping based on comparison selection of studies with proposed characterization framework
Result	Proposed Characterization Framework (Figure 9)

Table -3: Digital Libraries for Scholarly Search Databases for Primary Data Selection

(URLs)	
ACM	https://dl.acm.org/
IEEE	http://www.ieeexplore.ieee.org
Scopus	https://www.scopus.com/
Clarivate (W.O.S.)	https://clarivate.com/
Google Scholar	Google Scholar

Finding papers for this comprehensive review required searching through various digital libraries. This work prioritizes the broad worldview of the selected studies, as this is a review, not a technical paper. We have focused on analyzing the existing technologies through our proposed methods. Some of the sub-areas may be underrepresented in this study. However, as per our primary goal, this study will use search techniques to evaluate broad topics concerning data center energy-saving technologies in the Cloud, such as virtualization & consolidation, scheduling, and server management. (All of these techniques are classified and narrowed down in (section 4).

3.2 Data Duplication

Digital libraries are quite essential tools for study selection. However, their search results result in many unwanted effects; they help search the content with keywords, titles, abstracts, years, or other sets. Hence, it is common to end up with thousands of search results even if our queries are well-written. The selection of the data is excessively discussed in the next part.

3.3 Study Selection

The articles chosen for the primary studies are all subject-specific, i.e., "Energy Saving Technologies in Cloud." Thus, prime facets dealt with the selected items in the cloud computing domain. Key terminology, normally performed on the abstracts of peer-reviewed journals, is the foundation of a systematic mapping analysis. As a result, stories from newspapers, social media sites, and other outlets which fall into the gray study area are excluded.

3.4 Quality of Inclusion and Exclusion

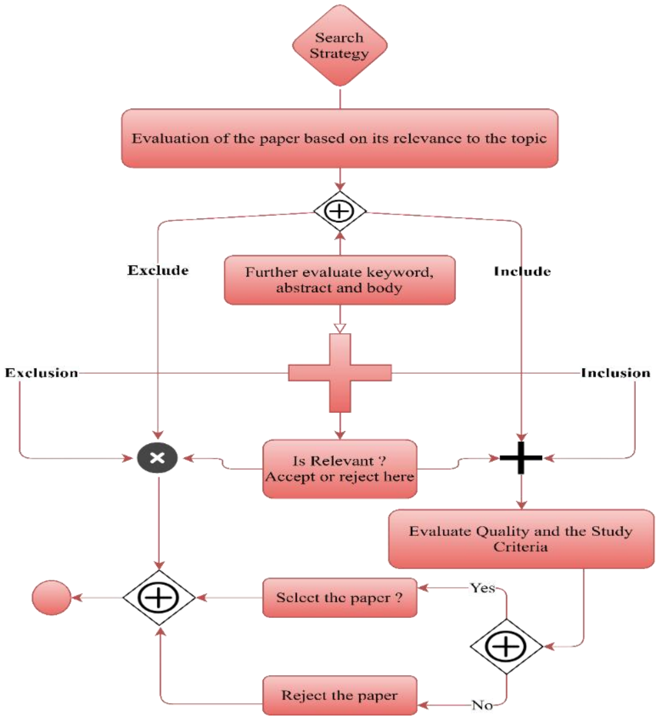
Our work's inclusion and exclusion mechanism are subject to rigorous extraction of primary data) Moreover, transformed into a secondary study to provide readers with a necessary literature background and peer-reviewed research articles. In addition to inclusion/exclusion criteria, a systematic extraction search model is applied to the process of search evacuation. Furthermore, we have used a study selection strategy based on [27] divided into two parts. First, we measured the primary date's significance in the first section by reading the description, abstract, and keywords; if that were not enough to make a meaningful conclusion, the full text was reviewed. Second, if the article is irrelevant, it will be excluded, and if the relevance is in question (due to issues among evaluators), a third party should be contacted.

Table -4: Inclusion and Exclusion Criteria

Inclusion	Our primary concern is to extensively highlight energy-efficient and energy-saving applications and technologies in Cloud data centers; this is a broad topic, as it includes the system side and networking. Our chief inclusion parameter lies within the application domain for this novel mapping and scoping study- Hence, we have explicitly double-checked the papers before including them in this work, see (Table - 7)
Exclusion	The literature not outlined in inclusion is excluded, including non-peer-reviewed material and misleading papers. Only English language papers are included.

Our selection strategy is based on the guidelines set by [27]. Our follow-up and cross-checks on the inclusion and exclusion criteria also prevent any unimportant data selection or pieces of gray literature from affecting our mapping analysis. Therefore, we carefully followed standard guidelines while selecting, extracting, and finally mapping.

Fig -3: The Study Selection Strategy [27]



3.5 Keywording

Keywording will help cut down on time it takes to create a classification scheme. This methodological step ensures that only suitable and sound papers are gathered. The keywording process necessitated a thorough examination of the abstracts to identify key terms, which can be helpful for the overall structure of the work. The keywords can further be divided into as:

Table -5: Keyword Selection Method

Title	Most important keyword for targeting primary data
Abstract	The abstract is extracted based on the relevance and soundness of the study
Classification Scheme	See Classification Framework (Figure – 4)
Mapping	Each step leading to an outcome

3.6 Classification Scheme

We have classified significant categories based on guidelines set by [28]. It sorts all the selected types based on a systematic approach for better categorization and visualization. In addition, we are using this technique as it applies to the work we have carried out and fallen into the cloud computing domain. We use the category as Contribution Type. The remaining classes are subject to technical, managerial, architectural, domain-centric, and generic domain aspects, particularly to answer questions from Q1 to Q5, emphasizing Q2 to Q4 as these comprise technical and quality concerns.

Fig -4 Categorization Scheme

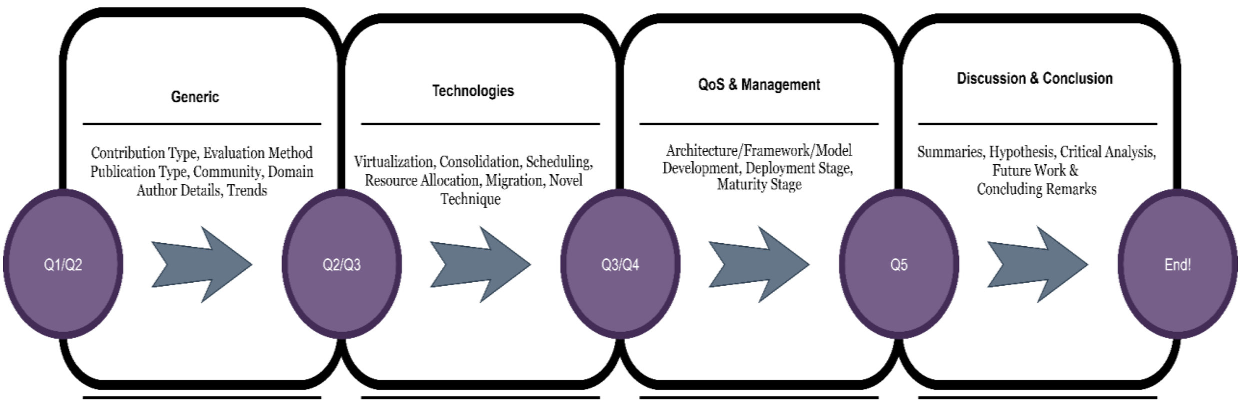


Table -6 Classification of Research Contribution

Validation Research	The methods tested were new but have yet to be implemented.
Evaluation Research	Methods are successfully implemented, and the efficacy of the process is evaluated.
Solution Based	An issue is addressed, and a resolution is suggested; the approach may well be new or a considerable improvement of an existing method.
Opinions & Discussions	Such papers share a person's opinion over whether a procedure is right or wrong or how things are done.
Experience	Experience articles describe what has already been done effectively and how it has been done. It must be the writer's own direct experience.

We have categorized our research scheme into four major components, each containing topics leading to various sub-topics. We align each with the proposed research question to have clarity of thought and clarity of facts. Our mapping study promises a systematic visualization, characterization, and categorization of the primary studies to synthesize a worldview and then narrow it into a personal research agenda, discussing each with facts accumulated from a deep analysis, validation, and prevention of authors' bias approach before representation.

3.7 Data Extraction & filtration of SMS

In addition, data extraction (Systematic Mapping) followed by extensive quality assessment and data filtration was also explained in the study selection. After that, we divided primary data into categories to represent the results. Furthermore, before analyzing the results, we have transformed our classification scheme into a conceptual framework (Fig - 5). In addition, the classification of the studies selected; was strengthened; some minor changes and revisions were necessary to sort data and keep its usefulness to the fullest. We also discussed how we gained our primary data in the previous sections. Before inclusion, we manually compiled a spreadsheet to input various digital sources while paying utmost attention to the keywording, primarily abstract and titles.

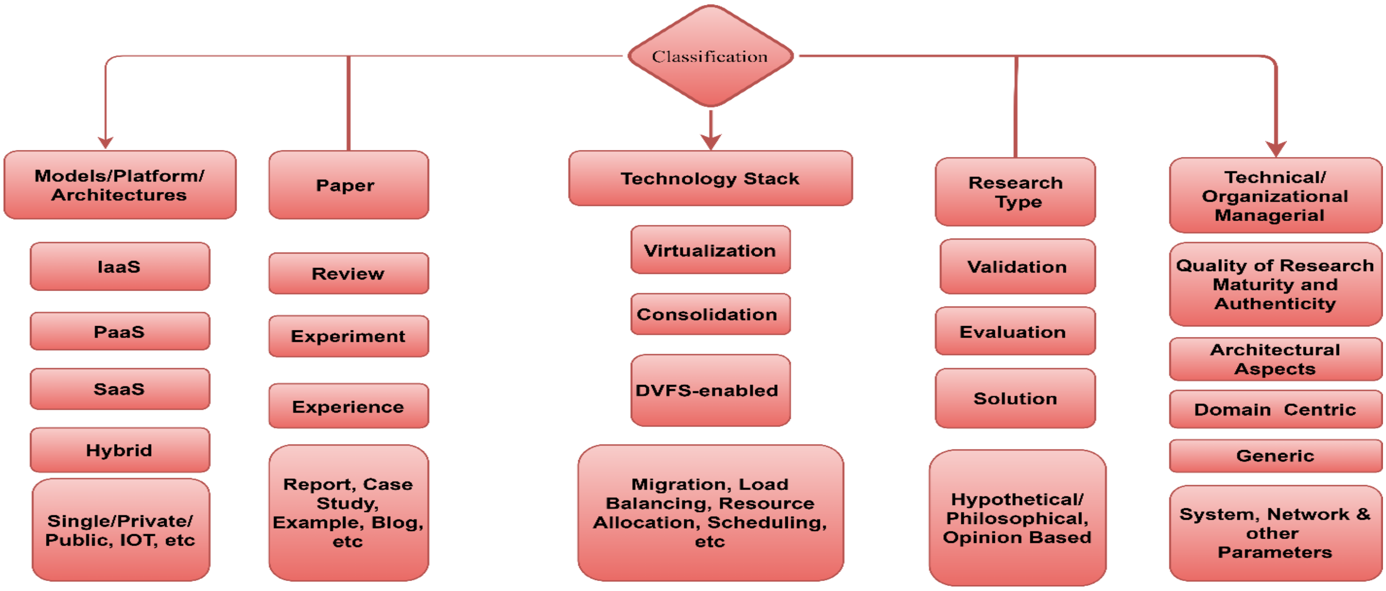
Moreover, we manually got rid of duplication of the source papers. We have used (Pivotal Tables) and some advanced Excel formulas to extract data into reader-friendly visualizations. Besides excel, we tried to use the R programming language and other bolometric libraries, which are familiar with the studies related to literature analysis. However, to our surprise, the credibility of such a process could be clearer. The authors must disclose how they merge two sources of different selections, for example, Web of Science and Scopus or IEE and Springer.

We, on the contrary, replied to the principles set by guidelines of [24-26] and updated guidelines of [29]. The author stressed the technology stack and generic results, such as venues, type of publication, frequency of each area of research, and others. We aim to comprehensively analyze the active and frequent technologies in the cloud domain. Hence, our job is to synthesize primary material to analyze useful concepts, the type of research contribution (Table- 7), and the venues and frequencies. Thus, we have divided two potions for the results: the first deals with the generic terms and the latter technical aspects. Because we can achieve the latter using any digital library online. Our focus on technology alone covers the most widely applied energy efficiencies and power management techniques, such as types of virtualizations and consolidation, scheduling, migration, resource allocation, provisioning, and dynamic scaling of the CPU power utilization. In addition, we kept a category for the state-of-the-art novel research other than these categories because we could not show substantial data in the table. Besides, primary data contain similarities, so we merged categories precisely to eliminate the information overflow. Moreover, categorizing terms such as contribution type and evaluation method has subtitles and subcategories similar for the management and administration. We have covered various subcategories, such as quality of service and service level agreement, execution, and maturity stages. Finally, there are system models, cloud delivery models, cloud settings, and other relevant sub-areas concerning architectural aspects.

3.8 Characterization Framework

In order to organize data collected into subject matter that is both reader-friendly and visually appealing, we have categorized our work into this characterization process (Table - 7). We were inspired by the system development life cycle principles used in software engineering project management for our proposed framework. It demonstrates the pertinent information and the classification standards.

Fig -5: Characterization Framework



Our characterization framework results from our characterization scheme, which resulted from the primary data selected. It is a conceptual framework that depicts major and minor categories, classes, and subclasses. There are four major components:

- Technological
- Research Contribution
- Management
- Platform/Architecture.

As discussed earlier, each step has subcategories elaborated in section V in-depth. This framework has been modified several times, as it is hard to classify a whole area of research. The second author and a guest researcher were approached to reflect and suggest preventing the author's preconception. However, the classification may underrate or overrate some research areas because of the vast selection of literature and broadness of the topic itself; however, as a big picture, it has been true to its very purpose, which is to address, energy-saving technologies for data centers in the Cloud.

4. RESULTS

We addressed the classification framework earlier; now, it is time to respond to the research questions from (section 3.1). This section will discuss those questions using graphics such as excel & origin charts, tables, and custom-made diagrams. We will begin with a general topic/concern attributed to research questions 1 and 2. After that, we will address technology-related concerns attributed to research questions 3 to 5. While we also identify the articles concerning publication format, platform, and technical engagement. In addition, we presented our findings' validity and implications for future studies. Starting from the first question, which is broad as it deals with the central topic of the proposed mapping study: "Data Center Energy-Saving Technologies in Cloud," we will see how active this research field is before classifying it into subcategories.

4.1 Overview of the Generic Questions

· What is the frequency of data center energy-saving research, and when did it become popular in the community?

· What active venues for energy-saving technologies in the Cloud, and which forums allow the most research in this category?

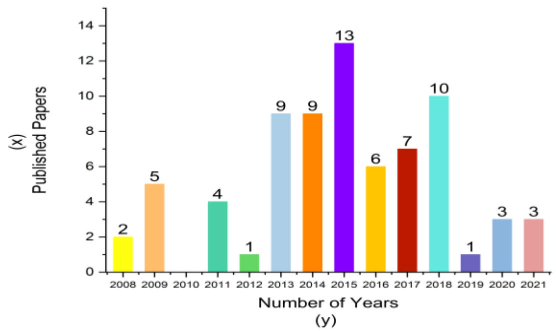
· Who are some of the most notable scholars, and what is their contribution concerning each technology?

· Which countries have produced the most research articles, and what is their frequency?

4.2 Frequency of Studies

The emphasis on energy saving by cloud computing platforms quickly grew as of 2008. This innovation was made possible by creating a market tightly connected to technology. Companies like Google, Amazon, and Facebook saw exponential growth. Such massive growth also brought with it real challenges. As a result, the investigation of energy-saving methods for data centers also emerged with a seemingly unrelated failure in 2012. However, the pace quickly picked up, particularly between 2013 and 2018. As a result, cloud service providers began constructing energy-efficient data centers. To create a narrative for additional analysis, we, therefore, considered the publications across several years. Our study selection is between May 2008 and May 2021. (the time of writing the article). By revealing how active this research topic has become over time, it will enable us to properly characterize it.

Fig -6 Yearly Publications Based on the Primary Study

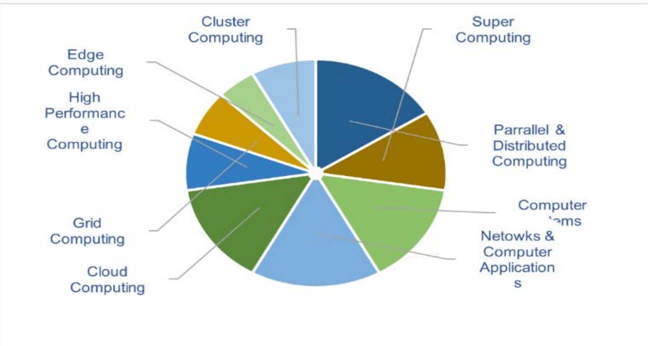


The total number of publications over time is depicted in Figure 6. This illustration was created using our main collection of data. Since energy concerns indicate using various IT components. It only applies to the software side of energy-saving techniques. We only used management and optimization techniques in software-related technologies. Such as virtualization and consolidation, green scheduling, and energy-efficient policies, models, and architectures, to justify our inclusivity. The figure above shows that the research area experiences certain variations each year because our primary data visualization uses data from articles pertaining to years. For instance, in 2010, there was no study. However, in 2015, there were 13 studies as opposed to just one study in 2019.

4.3 Overview of the Fora/ Communities

We have separated the subject into several research disciplines based on the major data analysis used for this study, including several engineering disciplines, computer science & systems, telecommunications, and software engineering, each representing a sub-discipline.

Fig -7 Primary Data Distribution Based on Community



Another key area to study is the active research categories and their respective disciplines: cloud computing, big data, edge computing, parallel and distributed computing, supercomputing, high-performance computing, grid computing, cluster computing, and network-related computer applications and systems. We can see the gradual increase of research activity on energy efficiency in nearly all cloud computing and computer science counterparts. We can also note that emerging areas like edge computing contribute immensely to the research before we categorize the papers based on our research questions, which will address in two ways: a. generic terms b. technical details. It is worth noting the following questions that directly relate to the general terms of the mapping study, for our goal is to address a broad perspective.

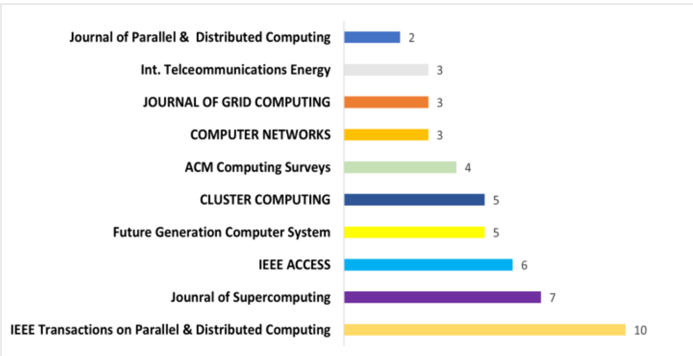
4.4 Overview of Publishing Sources & Active Researchers

The sources selected for our mapping study are peer-reviewed. Although we paid much attention to list out active journals, we have also considered reputable conferences, reports, and other distinguishable formats. We acknowledge and classify journals based on the frequency of articles on the topic related to our analysis. (Fig – 8) shows the topmost sources that published the most content on data center energy-efficient technologies and topic related.

4.5 Overview of Active Researchers & Countries

The categorization of research on the chosen topic based on nations and author preferences is another factor we noticed in our analysis. But we only classify certain things. We followed a strict normalizing procedure to eliminate redundant and irrelevant data. To accommodate only research that best connects to this topic and our purpose for a more vital subject-specific in-depth understanding of the region we focus on. Thus, we divided up research productivity and frequency by country.

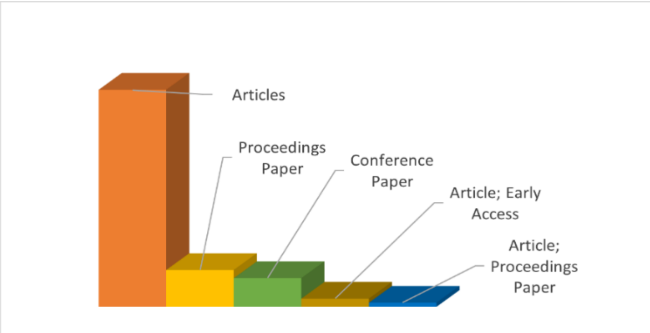
Fig -8: Topmost Journals That Publish Science Related to (Data Center Energy-Saving Technologies)



International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 09 Issue: 12 | Dec 2022 www.irjet.net p-ISSN: 2395-0072

Figure 9 shows various publishing formats. We divided the categories into articles, proceeding papers, conference papers, early access papers, and proceeding papers. Some of the most common formats of research publication, especially in reputable peer-reviewed journals in the domain. As we previously mentioned in community discussion, it is noticeable that parallel & distributed computing, grid computing, and others are some of the topmost active communities. Now we can see the journals based on those categories.

Fig-9: Distribution of Study by Publishing Formats



Over time, scientists may identify the leading countries producing a vast amount of quality research in energy-saving technologies in the Cloud. Figure 10 demonstrates the topmost active countries. China has produced the most significant number of papers in areas as diverse as cloud computing, big data, edge computing, and others. And the general reason behind this development is the number of affiliated research institutions in the country. Another reason is that the size of the economy is huge, which has enabled the researcher to conduct research.

Fig-10: Distribution of Study by Country

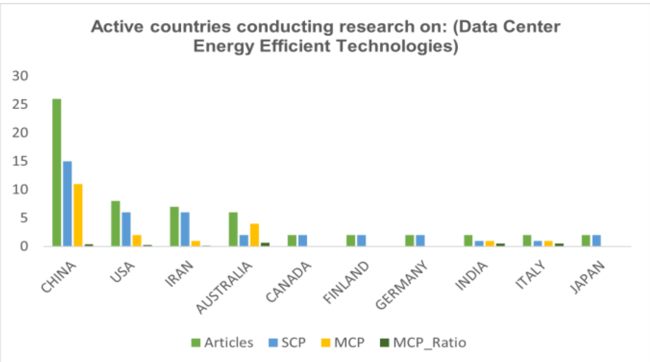
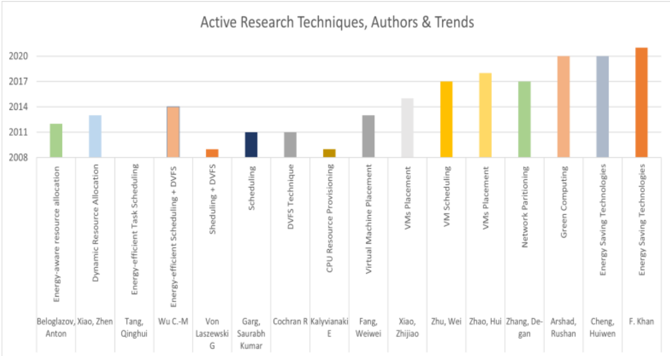


Figure 11 shows the specific technologies and writers who focused on those subjects. Besides using an automated R program, we also conducted a thorough manual analysis of the articles. Hence, studies are distributed by their top relevant authors.

Fig -11: Active Authors (State-of-the-art Techniques)



With 850 million smartphone devices as of the end of 2018, China had the highest number of smartphone subscribers among its counterparts, more nearly doubling the number of phone devices in the neighboring state (India) [30]. This aspect stimulates the demand for new technologies, including the emerging 5G technologies and a growing number of data centers. Besides, the need for cloud computing to boost customer quality of service is increasing. [31] Furthermore, the Chinese cellular market has actively promoted various technologies, namely semantic analysis, voice recognition, and computer vision and image processing techniques, which help boost further development of energy consumption; thus, the need for energy conversation also becomes necessary. It also holds the title of the highest-ranked country for the number of active cellular subscribers.[31] In addition, Cloud computing has received much attention from Chinese researchers, including energy efficiency in parallel and distributed computing. After China, the United States, Iran, Australia, Canada, Finland, Germany, India, and Japan has the most articles in the SCImago Country rank. [32].

We have chosen important publications that have helped develop dynamic methodologies and effective evaluation techniques. Additionally, we ensured that most of the articles we chose were peer-reviewed and had solid contributions, intricate design, real-world applicability, and modularity. In addition, we also created subcategories from the major study topics on cloud energy efficiency.

•For example, [33] has presented an energy-efficient resource allocation algorithm with the help of a consolidation-aware dynamic virtual machine placement technique while discussing opportunism and underlying challenges facing this field of study; their work provided a baseline and answered some questions for many researchers.

• On the other hand, [34] presented an energy-efficient resource management system in which they proposed a method that adaptively transfers virtual resources to physical resources. In addition, they have also used the skewness metric to integrate virtual machines with different resource requirements to adopt server capacity

• Another work [35] focused on a technique, MPIT-TA (Peak Inlet Temperature Minimization Problem). This method aims to improve the energy efficiency of a virtualized cloud environment and, ultimately, reduce SLA violation. Besides, they estimated that the proposed method could maximize the supply temperature and thus reduce air-conditioning energy requirements.

• The authors [36, 37] worked on resource utilization in a data center; their proposed method is based on Dynamic Voltage Frequency Scaling (DVFS) aware power management strategy. Moreover, they incorporated scheduling algorithms such as a DVFS-enabled approach to be of greater interest among many scientists. The number of DVFS based research papers has become popular, especially concerning green and energy-efficient scheduling algorithms and dynamic resource provisioning.

• This study [38] focused on implementing flexible resource provisioning and scheduling policies that take advantage of heterogeneity through multiple data centers. The proposed method aims to reduce energy conversation and costs.

• While work by these authors [39] implemented a method for the performance of multithreaded workloads on a multi-core CPU, their major proposal combines the DVFS technique into a scheduling algorithm by using real-time workload traces and benchmarking measurements for performance and validation of systems.

• Furthermore, the work done by [40] is based on virtualization techniques, especially Vm selection and placement problem; the authors' proposed technique deals with various aspects concerning a typical data center such as network performance, cost, and energy utilization, to address these concerns, authors worked on a virtual machine planner methodology to optimize energy.

• Moreover, [41] developed an algorithm by extending the evolutionary game theory to solve the virtual machine problem to optimize power. Their proposed model analyses the amount of power consumption during the process of adjusting a virtual machine.

• On the contrary, efforts were made to develop an Ant Colony optimization method to solve the VM placement problem (PPVMP) to obtain energy efficiency in virtualized cloud data centers.

• That leads to the work conducted by [42], which is based on a novel unequal clustering routing protocol that considers energy balancing based on network partition. In a nutshell, their effort was to create an energy-aware network clustering mechanism to solve energy conversation. Finally, the works of [6], [8] and, [43] have produced state-of-the-art surveys on emerging challenges facing green computing, data center energy efficiency, performance, architectural concerns, cost, and S.L.A. violation, to name a few.

4.4 Overview Technology Stacks

This section of the mapping study, related to research question number 3 to 5, will first evaluate each category's research contribution and evaluation methodologies. We have conducted our research on 74 studies. We have categorized each study by applying a conceptualized table (Table - 7), which extends our characterization framework. The method that we used is a unique approach to conducting a dynamic mapping study. Besides, the categorization we made to distinguish and exploit primary data is subject to a careful review of each selected study. Table 7 overviews all selected studies based on the most significant categories that range from technologies to methodologies. In addition, we focused on management concerns and quality of service, architectural aspects, and models. Multiple occurrences occur because many studies use similar techniques, and some apply a combination. However, we will classify each time into subcategories and sub-groupings for clarity and understanding. To begin with, we will first visualize results based on research question 3:

To subcategorize and evaluate research Q3

1. To compare the contribution type for understanding research direction in a particular area with an emphasis on a specific research methodology, technique, or strategy.

2. To reflect on the evaluation type and explore the various types to meet the requirements of our characterization framework?

4.4 Contribution Type

Based on our characterization framework, we have divided the selected studies for mapping into categories such as solution-based work, example, evaluation, experience, and validation. Multiple occurrences can be seen. However, our scheme and framework can be used as a reference model. As we have classified the research into various subcategories, it is also worth noting that we have aligned techniques within the contribution type result to have a clear view of which processes are commonly aligned with what popular methodologies. Figure 12 depicts the contribution type in which solution-based research is in high number. One reason is that the data center's centric research has gained huge momentum due to the large corporations' interest in the domain. Another reason is the nature of the research, for it involves controlled experiments of virtual machines to evaluate energy consumption, which is a solution-driven, experimental kind of research.

Fig -12: Study Division by Contribution Type

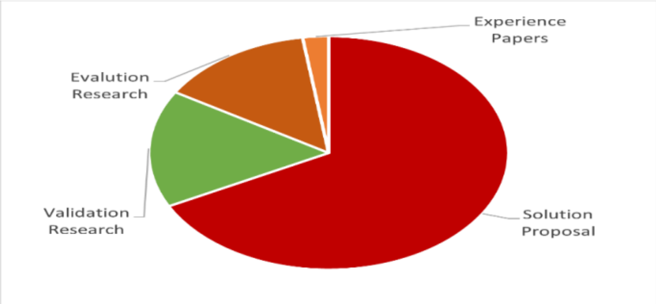
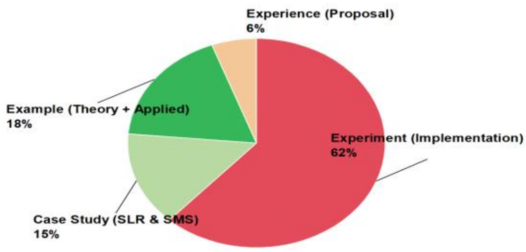


Figure 12 depicts most research based on solution proposal, followed by validation and evaluation. Each category shows a different method of research implementation based on 74 research articles. We have also demonstrated contribution type concerning evaluation technique, as we relied on manual readings and cross-checks. To reduce the author bias, we got the work cross-checked and proofread by guest reviewers.

4.5 Evaluation Method

The results of a few studies, including experimental research, example-based research, case studies, and experience-oriented research, are shown in Figure 13 in several categories. These are some of the required courses following the principles of software engineering. Like the solution suggestion, experiment-based research accounts for nearly 62% of the contribution type. Some articles utilized both the validation type of research and the evaluation type of research. Because of the technical specificity and the investigations' reliance on various approaches. Others also combined examples and experiences at the same time. Which is one reason some results can appear to be identical.

Fig – 13: Study Division by Evaluation Method



4.6 Technologies Stack

This section addresses the technology stacks based on previous terms such as contribution and evaluation methods. We focus on each technology now, and then we will present the results in the form of (Pie Charts %). The novelty of this work is that we classified each technology in this section. Since the technologies are well-integrated and rely on various methods, we will extract critical terms in the tables to have the best understanding.

Table -7 Conceptualization of Characterization Framework

Technology Stack	DVFS	VM(s)	[7] [8] [9, 10] [11, 12] [13] [14]	
	Resource Allocation & Provisioning		[15] [16] [17] [18]	
	Virtualization & Consolidation		[19] [20] [21] [22] [23] [24] [16] [25] [26] [27] [28] [29] [30] [31]	
	Scheduling		[32] [33] [16] [34] [35] [26] [36] [37] [11, 12, 38] [39] [7, 13, 40-42] [31, 43]	
	Dynamic Server Management		[44] [45] [46] [21] [22] [47] [48] [13, 49, 50] [25, 51] [43, 52, 53] [54]	
	Load Balancing		[55, 56]	
Methodology	Contribution Type	Solution Proposal	Adaptive + Generic	[57] [58] [59] [60] [61] [44] [45] [20] [46] [21] [22] [23] [47] [55] [7, 13, 50] [62] [33] [56] [63] [64] [16] [17] [34] [25] [51] [35] [26] [9] [36] [27, 65] [37] [28, 66, 67] [53, 68] [38] [10] [69, 70] [11] [7, 13, 39, 42, 71] [43] [31]

		Validation & Evaluation		[72] [21] [49] [50] [73] [74, 75]
		Review	SLR	[76] [77-79] [80-82] [83, 84] [85] [86] [87]
			SMS	
	Evaluation Method	Experiment		[57, 58] [59] [60] [61] [44] [45] [20] [46] [21] [22] [23] [47] [55] [7] [13] [62] [33] [56] [64] [16] [17] [34] [35, 51] [26] [9] [36] [27, 65] [37] [10] [69, 70] [11] [13] [7, 53, 71] [38, 67] [42, 66, 68] [43] [31]
		Case Study	Broad	[76] [73] [77, 78, 82, 84] [84] [81, 85] [86]
			Specific	[73] [79, 83] [80] [54, 87]
		Example	Theoretical	[72] [22] [48] [63] [73] [52] [28] [74, 75] [12] [41]
			Applied	[49, 50] [25]
Quality of Service & s Execution Management	Service Level Agreement	✔		[57] [20, 58] [22, 62] [33] [56] [64] [34, 36, 51, 65] [27, 67] [10, 66, 68, 70, 71] [69, 81] [39-42]
	Models Proposals	✔		[59] [72] [46] [22] [7, 49, 50] [63] [16] [25] [51, 74, 75] [80]
	Stage of Development	✔		[59] [45] [22] [47, 49, 62] [56] [63] [16] [25, 51] [28] [68, 84] [75] [69, 70] [13] [7, 40-42] [43]
	Deployment	✔		[57] [58, 59] [60] [61] [44] [45] [20] [46] [21] [22] [47] [55] [13, 50] [56] [64] [16] [25] [10, 27, 70] [75] [69]
	Workload or Benchmarks	✔		[58, 59] [59] [46, 64] [76] [34] [26] [26, 66]
	Architectural Concerns	Flexibility		[60] [61] [46] [55] [7, 33] [56] [63] [64] [34] [25] [36] [52] [28] [74] [66] [38, 53, 67, 85] [84] [78] [80] [70, 79, 81] [13] [86] [87]
		Modularity		[57] [58, 59] [61] [44] [45] [20] [46] [21] [22] [23] [47] [55] [7] [48] [13, 49, 62] [33] [63] [64] [16] [34] [25] [35] [26] [9] [37, 65] [27] [68] [10, 42, 54] [75] [41, 69, 71] [11] [12] [39] [7, 40] [43] [31] [31]
		Others		[59] [22] [47] [48, 49, 63] [28] [10] [75] [80]

The usage of virtualization and consolidation techniques are popular methods followed by scheduling and dynamic server management, including various hardware-centric processes, such as cooling mechanisms, energy storage systems, architectural aspects, and others. In addition, dynamic voltage and frequency scaling is becoming popular recently.

On the contrary, resource allocation and provisioning are also high, such as load balancing and virtual machine migration. We also focused on energy-efficient algorithms, models and frameworks, theoretical proposals, and alternative solutions-based approaches and containerization discussed in the results section in depth. Therefore, we have compared various techniques based on our categorization scheme and characterization framework in (Table - 7).

We have highlighted the results mapped from our primary data selection following the guidelines of our research questions. The table highlights the major components of the categorization. The reason behind this approach is to narrow down our mapping analysis and to enhance readability.

Fig – 14: Major Technologies

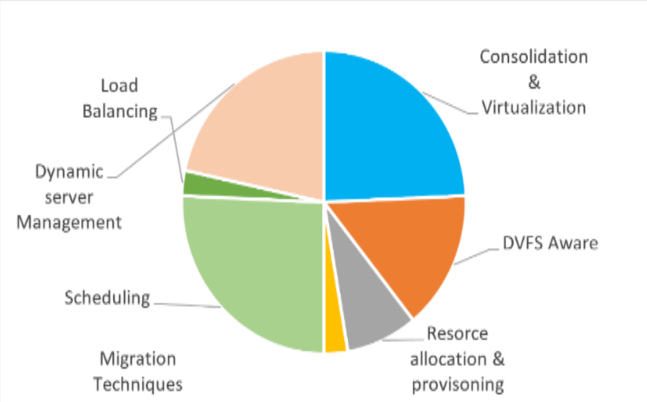


Fig – 15: Consolidation Classified

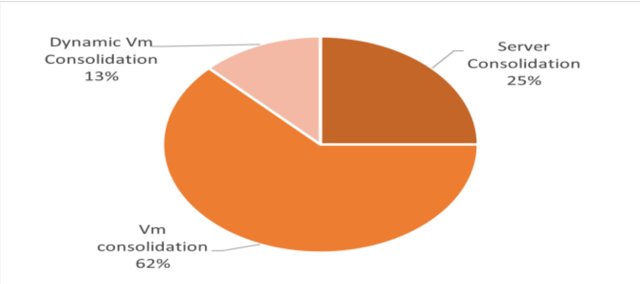
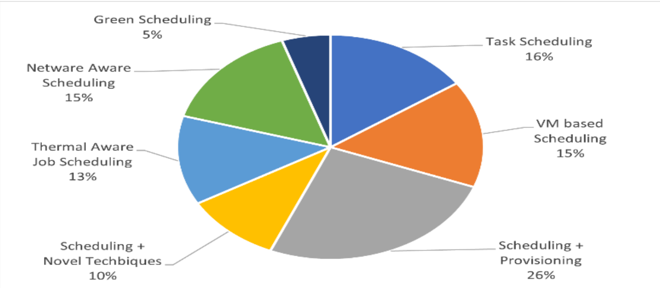


Fig – 16: Percentage of Scheduling-Based Research



We found scheduling based on resource allocation and provisioning methods to be in high number, approximately 26%, followed by network scheduling, task scheduling, virtual machine scheduling, thermal aware job scheduling. Our findings suggest that scheduling algorithms are prevalent in most studies.

In addition, figure 17 shows the number of DVFS-aware articles, the third most popular technique after scheduling and consolidation.

Fig – 17: Integration of DVFS with Scheduling

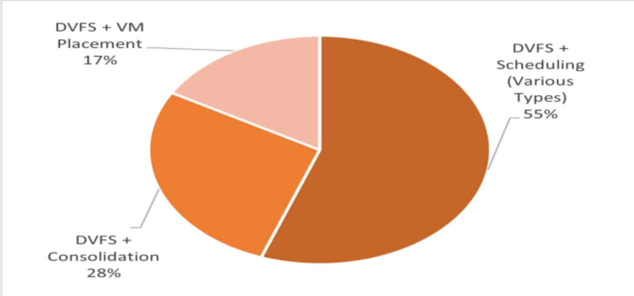


Table 8 is a collection of studies that used DVFS based scheduling algorithms. All the studies selected in the table are previously discussed in table 8, respectively, to prevent multiple occurrences; only references are included.

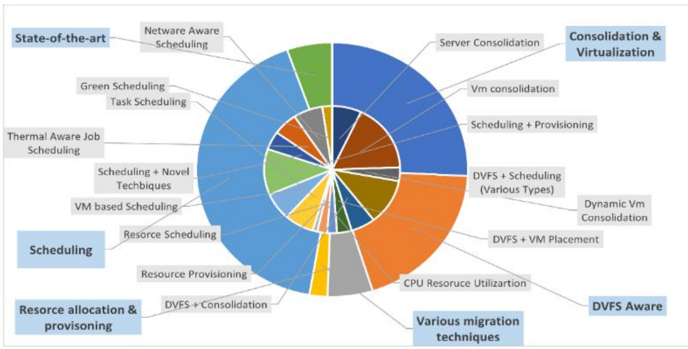
Table – 8: Technology Stack + DVFS (summarized)

Study	Technology Stack + DVFS (summarized)	Novel Model	Algorithm	Simulation	Real-Time Workload
[7]	Implementation of the energy consumption ratio (ECR) to assess the efficacy of specific wavelengths for executing a take so that it can adapt the power task scheduling problem of reducing the overall ECR	✔	Processor-level migration algorithm	✔	✔
[13]	DEWTS, an energy-saving scheduler based on a dynamic voltage/frequency scaling algorithm, is presented in this work	**X**	DEWTS	✔	✔
[35]	Dynamic Voltage and Frequency Scaling enabled Task Scheduling with a scheduling policy, design parameters, CPU Power Consumption for multi-core systems.	**X**	DVFS-enabled task scheduling algorithms	✔	✔
[89]	A frequency-aware and power strategy based on dynamic voltage and frequency have been proposed to lower power consumption.	✔	Heuristic scheduling algorithms	✔	✔

[10]	Power management strategy DVFS for implementing urgent CPU intensive bag of tasks	**X**	Cloud-Aware Energy-Efficient Scheduling	✔	**NA**
[88]	Green Energy Efficient Scheduling Algorithm (MMS-DVFS) based approach	**X**	DVFS-enabled green scheduling	✔	✔
[11]	DVFS enabled scheduling algorithm for VM placement, green scheduling approach.	**X**	Scheduling algorithm for DFVS-enabled clusters	✔	✔
[12]	Pack & Cap, robust method for improving the performance of multithreaded workloads on a multi-core CPU	**X**	DVFS-aware scheduling	**X**	✔
[13]	Based on DVFS-enabled workflow task scheduling (heuristic algorithm)	**X**	DEWTS	✔	**NA**
[7]	Experiment: Development of a new task model using power management approach (DVFS)	✔	MECRI, PTAB & Local Task Migration	✔	✔
[40]	Combination of Scheduling exploiting DVFS for tasks management in heterogeneous environment	✔	Energy Aware DVFS-enabled Scheduling	✔	**NA**
[42]	Job Scheduling into an energy objective model based on non-convex function	**X**	DVFS-enabled online adaptive and energy-aware scheduling algorithm	✔	✔

Figure 18 shows the frequency of recurrence of each term in the derived sentences from the samples. The key methods of the classification scheme are related to the term "occurrences":

Fig – 18: Percentage of All Major Technologies



4.8 Summary

First, we evaluated each technique individually. Second, we integrated most methods to find commonalities, interdependence, and alliance. Our effort has narrowed down specifics, characteristics, and underlying subcategories amongst the significant energy-saving technologies in the cloud data center.

4.8.1 Overview of Quality Concern (Q4)

We introduced quality of service criteria to address the power consumption issue concerning the following questions.

What are quality concerns in management and architectural elements?

- What are some of the tools, simulators, workloads, and frameworks within the selected studies?
- What are the deployment concerns?

In this section, we focus on the management and architectural concerns starting with the quality of service parameters with respect to selected studies (Table - 10). We classified primary attributes from studies that rely on simulation and use real-time workload traces or benchmarking suites (Table- 11). Aside, we classified models and frameworks that report SLA violations. We also address architectural concerns such as flexibility, modularity, integration, and adaptiveness, listed in (Fig - 19). Besides, we calculated the percentage of simulation-based research. The most popular simulation tool is CloudSim, especially concerning the implementation of virtualization/consolidation, scheduling, DVFS, and migration-based energy-efficient algorithms, as these techniques exploit the services and functionality by extending the built-in classes of the CloudSim simulator. While customized simulators either use a platform or integrate their model independently of any outsourced simulator, primarily using their model and programming technique. Then, we also focused on performance

parameters such as cost, and environmental concern, at the same time. In this regard, most studies used green scheduling algorithms for resource management and distribution. Moreover, we included provisioning based on the extracted keywords within the studies.

Table – 10: Quality of Service and Management

Count of QoS		
SLA Parameters	Reliability + Performance + Design + Modifiability	30
Infrastructure Parameters	Datasets + workloads + benchmarks + clusters	23
System Parameters	Disk + Storage + Cooling Systems + Cloud Infrastructures	34
Simulation	CloudSim + Green Cloud + Customized & Commercial	36
Performance	Cost + Time + Resources Management + Distribution + Provisioning	62

QoS helps understand the key findings, and creates a learning curve, builds a narrative by pointing out the most significant properties of the selected studies. Although many studies claim to reduce SLA violations, the rates are not high enough, so there is a lot of research gap in this very area. SLA breaches impact shareholders and users. We observed that clients, cloud providers, and infrastructure providers affected stakeholders based on the existing literature or that not recognizing SLA violations significantly affects users.

And service providers profit from implementing virtualization and consolidation techniques, as articles that discussed these methods reported (limited scope of enhancement in cost reduction, efficient resource consumption, and management). Moreover, we derived keywords from our keyword extraction strategy. According to our findings, the respective performance features in most of the studies stressed the keywords and attributes such as resource management, cost, time, distribution, and provisioning.

Table – 11: Workload and Benchmarking

Real-time Workload Traces & Benchmark Suites Count	
Planet Labs	7
Google datacenter traces	2
HiBench benchmark suite	1
K means Clustering	1
Google Cluster Data Datasets	1
SPECpower benchmark.	3
PARSEC parallel benchmark suite	1
Rubis benchmark website	1

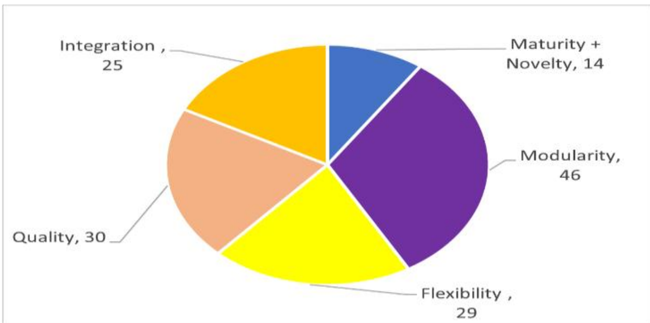
Table 11 depicts some of the most used real-time workload traces. We can see the Planet Lab is high, and all others are deficient. We noticed that the articles that used simulation-based methods to perform their operations heavily relied on real-time workload traces, especially works that extended the CloudSim simulator.

Table – 12: Major Frameworks with Notable Studies

Notable mentions of Frameworks	
Lyapunov optimization framework	[63]
Profiling-based server consolidation framework	[20]
Data Center-wide Energy-Efficient Resource Scheduling framework (DCEERS)	[33]
Architectural Framework	[90]
Framework for "energy efficiency of MapReduce Applications	[34]

Another quality concern within the management category is the framework and models. It implies that a particular solution from one of the selected studies in this paper relied on one of the listed frameworks (Table 12). In addition, figure 19 depicts the architectural concerns found in our primary study selection.

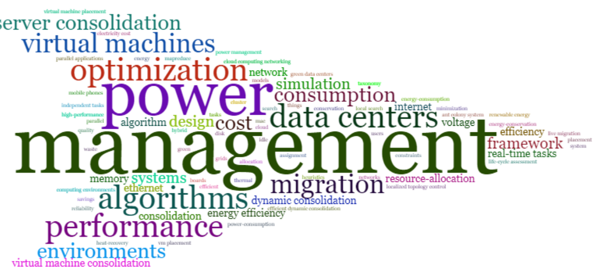
Fig – 19: Architectural Concerns (Based on Primary Study's Keywords Extraction Analysis)



Based on our conceptual table, which extended our characterization framework, we specified the generalized terminologies to reflect on the quality of service of the studies from the perspective of the architectural characteristics, which accounts for significant attributes such as flexibility, modularity, integration, maturity, and quality aspects. Modularity stands out based on the result (Fig - 19), meaning that the studies primarily used solution-based research for their evaluation methods.

Moreover, we extended our results in each section to classify major categories into various subcategories. Our classification is subject to our proposed scheme. Figure 20 showcases the most used keywords derived from the complete text, mainly focusing on the abstracts of all the selected papers. It unfolds that we systematically discussed the motive of this study while implementing our classification criteria.

Fig – 20: Extracted Keywords from all Studies



Even though we have followed a rigorous method to implement our study, we have also mentioned threats to the validity within each section. Besides, we systematically avoided multiple occurrences of similar terms as two authors searched for the keywords to reduce the similarities and enhance the scientific accuracy. In addition, we created a workbook to record our keepings between two authors to prevent any bias. Additional guest authors evaluated and proofread the selection to enhance the quality before merging the table into a single source document. In addition, we discarded grey literature, non-peer-reviewed articles, and articles whose content mismatched their abstracts. Besides, we applied the search quality mechanism, inclusion, exclusion, and PICO criteria (see Section 3). We filtered primary studies based on the study selection strategy (Fig - 7).

4.9 Simulation

Experimentation with new methods in a real-world Cloud environment is complex since some tests affect the end-user quality of service. Testing in the real-world environment evaluates a system's performance, but it is expensive and time-consuming [91]. In addition, the consumer does not have access to all the components of the device that need to be evaluated. Simulation distinguishes quality problems and enables the ideal issue to be concentrated [92]. Cloud simulators provide a stable, cost-efficient, and scalable environment where the testing community can reproduce, repeat, and validate experiments. They allow users to monitor all layers of the cloud system and the configuration of physical resources, the topology of the infrastructure, the middleware code platform, the services of the cloud application, and the actions of user workload [92]. There are a few noteworthy modelling tools related to cloud computing to evaluate and monitor the effectiveness of most energy-efficient policies and algorithms. CloudSim toolkit is perhaps the most famous simulation software for resource scheduling and planning. It is used to measure different QoS metrics, such as makespan time, power consumption, accessibility, dependability, energy usage, using sustainable classes based on algorithm criteria. See the following most popular simulation tools available for modelling Cloud computing resources.

4.10 Progress in the field

Virtual machines have several issues, including server failures due to consolidation, network congestion, and sprawl waste. Performance, management, and cost are all major problems as well. Researchers are always looking for more efficient and powerful technologies. Containerization is a potential study area that has recently attracted much scholarly interest. Since it allows applications to run in distinct user areas (containers) that all share the same OS kernel. As it helps make the formation and deployment of applications quicker and much safer. Furthermore, it is not unusual for problems when code executes using conventional methods in one computer network and then switches to a new console. Containerization captures the entire source code and all of its interconnections and file system. Therefore, lightweight virtualization technology rapidly increased, and today's most predominant idea is containerization (or containerization). In the maritime industry, container technology refers to processing an application to separate its dependencies. Microsoft Azure, Amazon Web Services (AWS), and Google Cloud Platform have embraced container technology and widely deployed it in hyper-scale data centers. When contrasted with virtual machines, containers can share the OS kernel and reduce the need to identify an operating system for each implementation. Since containers get much smaller than a VM and require very little startup time, they could operate many more occurrences on a single server. There is indeed a reduction in energy consumption for the same number of applicants. For those that want to put docker containers into exercise, here are among the most likely choices: KataContainer (by OpenStack) by IBM, Linux LXC by Docker, FireCracker, and Shifter (by IBM) (by Amazon). Unikernel is an intelligent new technology that has the potential to improve the advantages of lightweight virtualization. Unikernel, also recognized as container 2.0, takes simplicity to the next level by giving only the applications required to run the selected software. Regarding technicalities, Unikernel relies on specialty implementations to integrate software and support OS features during the execution stage rather than at runtime. A standalone executable image contains everything that the implementation gets to execute the consequence of this procedure. For building virtual machines, energy saving is a primary concern. When we reduce the complexity of virtual machines, the servers conceivably become more efficient. There is a price for additional benefits, and it weakens the price segregation among applications, which might lead to quality and energy assertion troubles.

4.11 Renewable Provisioning based on Load Scheduling Techniques

Renewable energy has a charm, and it is no doubt that tech-savvy companies want to invest in a potential business opportunity while enhancing their performance and raising standards. It is a fact that renewable energy is not an alternative now; the overall utilization ratio is lower than 1%. However, it can change in the future because the trends support this argument. Facebook promised to use only renewable power in 2011. In 2012, Google (the world's biggest tech giant corporate buyer of renewable energy) and Apple accompanied. Almost 20 Internet providers have completed the same as in 2017. Based on the facts, big companies only seriously considered renewable energy alternatives recently. However, the trend has shifted already. Enormous investment is being made by the top-most tech giants, namely Amazon, Apple, and Google. Such a trend is healthy for academia and industry. A smart and innovative AI-driven approach to optimize power-intensive infrastructure is a way forward to green and clean data centers. Following are some outstanding studies discussing, reviewing, and assessing probable solutions to bridge the gap of energy crisis and will use renewable energy as a source of power data centers: [93-98]

4.12 Optimization of Energy With Artificial Intelligence

Datacenter server management was no exception to the general rule in the past couple of years: an empirical study and domain expertise are the go-to methods to determine ideal setups and strategic goals. Recent technologies such as artificial intelligence (AI) have gained widespread attention. Artificial intelligence alternatives have succeeded in various real-time computational applications that people consider complex and puzzling human jobs. Dynamic power management of servers in data centers is a complicated and cost-heavy process, as discussed throughout the paper. DVFS-based AI methods can work in some ways, but we still need a set of tried and tested benchmarks to follow. Most independent studies need to provide more evidence on how to opt for the exact configuration when dealing with DVFS. However, few studies showed promising results and supported the theory for real-time AI-driven power-efficient data center resource management. One such instance is.[99] They proposed a deep enforcement learning method for optimizing DC cooling control in the data center. According to the modeling system's results, the suggested CCA could save up to 11% on cooling costs compared to a manually designed benchmark controller.[99] Another study [100] showed a practical example of AI-driven optimization to reduce energy consumption in a data center. For example, computer-controlled cloud provisioning on AWS using deep reinforcement learning recommends using Deep Reinforcement Learning (RL) to digitize cluster scale-out/in as an alternative approach. Using Q learning, they designed a good cluster controller that discovers how and where to take good actions in a specified Q-state, considering the consequent condition and recompense. These are some of the useful studies in the light of AI-enabled power-efficient and cost-effective data centers:[101-103]

5.0 Open Issues & Discussion

Because the cloud is a business strategy, we must regard users' focus throughout implementation. Energy efficiency is a complicated environmental problem; thus, a substantial study is necessary for energy-based resource provisioning. Cloud service providers focus on their business model; there ought also to be an emphasis on the users' quality of service parameters, energy efficiency, and usability goals. Besides, some characteristics in the implementation process of Cloud performance through modeling and simulation ought to be performed in a thorough and controlled setup. Such as identifying a threshold level, VM migration, tracking CPU or disk usage, and task migration.

Moreover, even though determining the incoming load in a private cloud is tough and more efficacious, researchers ought to analyze and establish workload prediction technologies by utilizing (deep learning, artificial intelligence, and machine learning techniques). In addition, energy efficiency is a complicated environmental problem; thus, a substantial study is necessary for energy-based resource provisioning.

Furthermore, researchers must ensure security concerns, such as network sensitivity and vulnerabilities, to prevent data loss and information leakage. For robust resource management, some advanced novel models must be incorporated and aligned with current dynamic task scheduling algorithms. CSPs and independent academic researchers can pay more attention to multi-aim scheduling algorithms that include cost-aware models, and heuristic and meta-heuristic technologies that satisfy usability goals should be considered. Moreso, CSPs and independent academic researchers can pay more attention to multi-aim scheduling algorithms that include cost-aware models, and heuristic and meta-heuristic technologies that satisfy usability goals should be considered. Ultimately, we can focus more on reducing SLA violations, especially while provisioning resources.

5.1 Threats to Validity

Our topic centers on energy-efficient data center technologies. We researched multiple research significance viewpoints. To prevent the bias of generalization before the very selection of the primary studies, we focused on the software engineering side of the energy-saving technologies; we shortlisted only technologies that authors experimented with the help of software-centric tools since the involvement of the system and networking is quite repetitive and at various stages. Hence, we mitigated multiple occurrences, which we either discussed in relevance or discarded in case of repetitiveness. Our prime goal is to discuss, evaluate and reflect on virtualization and consolidation technologies, scheduling, resource allocation & provisioning, and dynamic server management. However, server management has many research significances to undertake from the hardware perspective to the software. However, we only manifested the terms and results suitable for the software-centric approaches, such as server consolidation techniques.

5.2 Conclusion

Green data centers involve various power management approaches. Optimization of energy has the most significant factor in achieving the goal of green energy. In this paper, we assessed data center energy-saving technologies and software aspects of energy-efficient techniques. We focused primarily on virtualization and consolidation approaches while discussing and evaluating various scheduling algorithms, state-of-the-research, dynamic server management, resource allocation and provisioning, migration, and load balancing, to name a few. In addition, we stressed energy-efficient software-centric technologies in the Cloud, which helps boost the performance of the data centers and helps reduce cost for the Cloud Service Providers (CSP.). We applied a systematic mapping study on selected primary studies (74) publications. Besides, our approach comprises a series of systematic steps: a quality mechanism of selection of the studies including manual, snowballing as well as database, a search strategy into the most reputed digital libraries, an extraction method, a characterization framework to record relevant data for classification, a conceptual scheme to collect primary data. Moreover, the results are distributed based on a) generic Systematic Literature Review (SLR), b) Dynamic Systematic Mapping Study (SMS). In addition, we classified the contribution type of each study and the evaluation method. The relationship between techniques and processes by listing out the primary implementation flow of almost all effective ways by extracting essential contents from the primary data. We have found that nearly 60% of the study is solution-based, which involves experimentation. Model development, integration, and customization of techniques are popular. Since 2018, containerization-based research is emerging. They shifted more and more focus on the environmental aspects of Cloud Data Centers. In brief, virtualization and consolidation-based analysis appear to be one of the most implemented, especially concerning dynamic server consolidation. Whereas workload scheduling and scheduling algorithms also seemed promising in their regard, especially DVFS-enabled energy-efficient algorithms with customized provisioning policies for resource allocation and distribution. Besides, CloudSim based simulation is also popular in many types of research that we focused on in this paper. In which, authors are applying DVFS and Scheduling techniques for power management. Apart from this, load balancing and live migration techniques have seen a drastic decline in popularity. Techniques such as containerization, AI-based energy efficiency methods, and edge computing concepts are getting popular. Our graphs also visualized QoS, SLA, development, deployment, architectural concerns, workload, and benchmarking. In addition, we presented our results to focus on the models, frameworks, tools, platforms, performance, and other technical aspects. In addition, we reflected a learning curve altogether in a systematic, reader-friendly visualization. In sum, we have evaluated major data center energy-saving technologies concerning virtualization & consolidation-based software-centric methods in this work. Hence, our paper provides a roadmap and can be used as a reference model for many scientists who want to study energy-efficient techniques in cloud data centers.

REFERENCES

1. Cheng, H., et al., *A survey of energy-saving technologies in cloud data centers.* The Journal of Supercomputing, 2021.

2. Kaur, T. and I. Chana, *Energy Efficiency Techniques in Cloud Computing: A Survey and Taxonomy.* 2015. **48**(2 %J ACM Comput. Surv.): p. Article 22.

3. Khan, F., et al., *Software techniques for making cloud data centers energy-efficient: a systematic mapping study*. 2020 46th Euromicro Conference on Software Engineering and Advanced Applications. 2020. 479-86.

4. Abadi, R.M.B., A.M. Rahmani, and S.H. Alizadeh, *Server consolidation techniques in virtualized data centers of cloud environments: A systematic literature review.* Software-Practice & Experience, 2018. **48**(9): p. 1688-1726.

5. Mastelic, T., et al., *Cloud Computing: Survey on Energy Efficiency.* 2014. **47**(2 %J ACM Comput. Surv.): p. Article 33.

6. Kitchenham, B., et al., *Systematic literature reviews in software engineering-A systematic literature review.* Information and Software Technology, 2009. **51**: p. 7-15.

7. Wang, S.Y., et al., *A DVFS Based Energy-Efficient Tasks Scheduling in a Data Center.* Ieee Access, 2017. **5**: p. 13090-13102.

8. !!! INVALID CITATION !!! [8, 9].

9. Zhou, Z., et al., *Minimizing SLA violation and power consumption in Cloud data centers using adaptive energy-aware algorithms.* Future Generation Computer Systems, 2018. **86**: p. 836-850.

10. Calheiros, R.N., R. Buyya, and Ieee, *Energy-Efficient Scheduling of Urgent Bag-of-Tasks Applications in Clouds through DVFS*, in *2014 Ieee 6th International Conference on Cloud Computing Technology and Science*. 2014. p. 342-349.

11. von Laszewski, G., et al., *Power-Aware Scheduling of Virtual Machines in DVFS-enabled Clusters*. 2009 Ieee International Conference on Cluster Computing and Workshops. 2009. 144-153.

12. Cochran, R., et al., *Pack & Cap: Adaptive DVFS and Thread Packing Under Power Caps*, in *Proceedings of the 2011 44th Annual Ieee/Acm International Symposium on Microarchitecture*. 2011. p. 175-185.

13. Tang, Z., et al., *An Energy-Efficient Task Scheduling Algorithm in DVFS-enabled Cloud Environment.* Journal of Grid Computing, 2016. **14**(1): p. 55-74.

14. !!! INVALID CITATION !!! [8, 15-17].

15. !!! INVALID CITATION !!! [17, 18].

16. Mukherjee, T., et al., *Spatio-temporal thermal-aware job scheduling to minimize energy consumption in virtualized heterogeneous data centers.* Computer Networks, 2009. **53**(17): p. 2888-2904.

17. Mishra, S.K., et al., *An adaptive task allocation technique for green cloud computing.* Journal of Supercomputing, 2018. **74**(1): p. 370-385.

18. !!! INVALID CITATION !!! [21, 22].

19. !!! INVALID CITATION !!! [27, 28].

20. Ye, K.J., et al., *Profiling-Based Workload Consolidation and Migration in Virtualized Data Centers.* Ieee Transactions on Parallel and Distributed Systems, 2015. **26**(3): p. 878-890.

21. Yang, T., et al., *Energy-Efficient Data Center Networks Planning with Virtual Machine Placement and Traffic Configuration*, in *2014 Ieee 6th International Conference on Cloud Computing Technology and Science*. 2014. p. 284-291.

22. Yadav, R., et al., *Adaptive Energy-Aware Algorithms for Minimizing Energy Consumption and SLA Violation in Cloud Computing.* Ieee Access, 2018. **6**: p. 55923-55936.

23. Xiao, Z.J., et al., *A solution of dynamic VMs placement problem for energy consumption optimization based on evolutionary game theory.* Journal of Systems and Software, 2015. **101**: p. 260-272.

24. !!! INVALID CITATION !!! [33, 34].

25. Lovasz, G., F. Niedermeier, and H. de Meer, *Performance tradeoffs of energy-aware virtual machine consolidation.* Cluster Computing-the Journal of Networks Software Tools and Applications, 2013. **16**(3): p. 481-496.

26. Li, X., et al., *Holistic Virtual Machine Scheduling in Cloud Datacenters towards Minimizing Total Energy.* Ieee Transactions on Parallel and Distributed Systems, 2018. **29**(6): p. 1317-1331.

27. Horri, A., M.S. Mozafari, and G. Dastghaibyfard, *Novel resource allocation algorithms to performance and energy efficiency in cloud computing.* Journal of Supercomputing, 2014. **69**(3): p. 1445-1461.

28. Ghribi, C., M. Hadji, and D. Zeghlache, *Energy Efficient VM Scheduling for Cloud Data Centers: Exact allocation and migration algorithms*, in *Proceedings of the 2013 13th Ieee/Acm International Symposium on Cluster, Cloud and Grid Computing*, P. Balaji, D. Epema, and T. Fahringer, Editors. 2013. p. 671-678.

29. !!! INVALID CITATION !!! [38, 39].

30. !!! INVALID CITATION !!! [39-41].

31. Abohamama, A.S. and E. Hamouda, *A hybrid energy-Aware virtual machine placement algorithm for cloud environments.* Expert Systems with Applications, 2020. **150**.

32. !!! INVALID CITATION !!! [41, 42].

33. Shuja, J., et al., *Data center energy efficient resource scheduling.* Cluster Computing-the Journal of Networks Software Tools and Applications, 2014. **17**(4): p. 1265-1277.

34. Mashayekhy, L., et al., *Energy-Aware Scheduling of MapReduce Jobs for Big Data Applications.* Ieee Transactions on Parallel and Distributed Systems, 2015. **26**(10): p. 2720-2733.

35. Lin, C.C., et al., *Energy-efficient task scheduling for multi-core platforms with per-core DVFS.* Journal of Parallel and Distributed Computing, 2015. **86**: p. 71-81.

36. Kliazovich, D., P. Bouvry, and S.U. Khan, *DENS: data center energy-efficient network-aware scheduling.* Cluster Computing-the Journal of Networks Software Tools and Applications, 2013. **16**(1): p. 65-75.

37. Gu, C.L., et al., *Greening cloud data centers in an economical way by energy trading with power grid.* Future Generation Computer Systems-the International Journal of Escience, 2018. **78**: p. 89-101.

38. Ebrahimirad, V., M. Goudarzi, and A. Rajabi, *Energy-Aware Scheduling for Precedence-Constrained Parallel Virtual Machines in Virtualized Data Centers.* Journal of Grid Computing, 2015. **13**(2): p. 233-253.

39. Kalyvianaki, E., et al., *Self-Adaptive and Self-Configured CPU Resource Provisioning for Virtualized Servers Using Kalman Filters*. Acm/Ieee Sixth International Conference on Autonomic Computing and Communications. 2009. 117-126.

40. Safari, M. and R. Khorsand, *Energy-aware scheduling algorithm for time-constrained workflow tasks in DVFS-enabled cloud environment.* Simulation Modelling Practice and Theory, 2018. **87**: p. 311-326.

41. Arroba, P., et al., *DVFS-Aware Consolidation for Energy-Efficient Clouds*, in *2015 International Conference on Parallel Architecture and Compilation*. 2015. p. 494-495.

42. Shojafar, M., et al., *An Energy-aware Scheduling Algorithm in DVFS-enabled Networked Data Centers*. Proceedings of the 6th International Conference on Cloud Computing and Services Science, Vol 2, ed. J. Cardoso, et al. 2016. 387-397.

43. MirhoseiniNejad, S., G. Badawy, and D.G. Down, *Holistic thermal-aware workload management and infrastructure control for heterogeneous data centers using machine learning.* Future Generation Computer Systems-the International Journal of Escience, 2021. **118**: p. 208-218.

44. Yu, L., T. Jiang, and Y. Cao, *Energy Cost Minimization for Distributed Internet Data Centers in Smart Microgrids Considering Power Outages.* Ieee Transactions on Parallel and Distributed Systems, 2015. **26**(1): p. 120-130.

45. Yoshii, A., et al. *Development of a rack-type air-conditioner for improving energy saving in a data center*. in *INTELEC 2009 - 31st International Telecommunications Energy Conference*. 2009.

46. Yao, J.G., et al., *Adaptive Power Management through Thermal Aware Workload Balancing in Internet Data Centers.* Ieee Transactions on Parallel and Distributed Systems, 2015. **26**(9): p. 2400-2409.

47. Xiao, Z., W.J. Song, and Q. Chen, *Dynamic Resource Allocation Using Virtual Machines for Cloud Computing Environment.* Ieee Transactions on Parallel and Distributed Systems, 2013. **24**(6): p. 1107-1117.

48. Tsai, C.H., et al., *Online Web cluster capacity estimation and its application to energy conservation.* Ieee Transactions on Parallel and Distributed Systems, 2007. **18**(7): p. 932-945.

49. Tang, Q.H., S.K.S. Gupta, and G. Varsamopoulos, *Energy-efficient thermal-aware task scheduling for homogeneous high-performance computing data centers: A cyber-physical approach.* Ieee Transactions on Parallel and Distributed Systems, 2008. **19**(11): p. 1458-1472.

50. Takahashi, M., et al., *Aisle-capping Method for Airflow Design in Data Centers*, in *Intelec 08 - 30th International Telecommunications Energy, Vols 1 and 2*. 2008. p. 202-208.

51. Long, S.Q., Y.L. Zhao, and W. Chen, *A three-phase energy-saving strategy for cloud storage systems.* Journal of Systems and Software, 2014. **87**: p. 38-47.

52. Ikebe, H., et al., *Green energy for telecommunications*, in *Intelec 07 - 29th International Telecommunications Energy Conference, Vols 1 and 2*. 2007. p. 750-755.

53. Fang, W.W., et al., *VMPlanner: Optimizing virtual machine placement and traffic flow routing to*

reduce network power costs in cloud data centers. Computer Networks, 2013. **57**(1): p. 179-196.

54. Revesz, A., et al., *Developing novel 5th generation district energy networks.* Energy, 2020. **201**.

55. Xiao, X., et al., *A Workload-Aware VM Consolidation Method Based on Coalitional Game for Energy-Saving in Cloud.* Ieee Access, 2019. **7**: p. 80421-80430.

56. Rossi, F.D., et al., *E-eco: Performance-aware energy-efficient cloud data center orchestration.* Journal of Network and Computer Applications, 2017. **78**: p. 83-96.

57. Zhu, W., Y. Zhuang, and L. Zhang, *A three-dimensional virtual resource scheduling method for energy saving in cloud computing.* Future Generation Computer Systems-the International Journal of Escience, 2017. **69**: p. 66-74.

58. Zhou, Z., et al., *Minimizing SLA violation and power consumption in Cloud data centers using adaptive energy-aware algorithms.* Future Generation Computer Systems-the International Journal of Escience, 2018. **86**: p. 836-850.

59. Zhou, Z., et al., *Fine-Grained Energy Consumption Model of Servers Based on Task Characteristics in Cloud Data Center.* Ieee Access, 2018. **6**: p. 27080-27090.

60. Zhao, H., et al., *Power-Aware and Performance-Guaranteed Virtual Machine Placement in the Cloud.* Ieee Transactions on Parallel and Distributed Systems, 2018. **29**(6): p. 1385-1400.

61. Zhang, D.-g., et al., *Novel unequal clustering routing protocol considering energy balancing based on network partition & distance for mobile education.* Journal of Network and Computer Applications, 2017. **88**: p. 1-9.

62. Singh, S., et al., *SOCCER: Self-Optimization of Energy-efficient Cloud Resources.* Cluster Computing-the Journal of Networks Software Tools and Applications, 2016. **19**(4): p. 1787-1800.

63. Qin, Y., et al., *Joint energy optimization on the server and network sides for geo-distributed data centers.* Journal of Supercomputing.

64. Hieu, N.T., M. Di Francesco, and A. Yla-Jaaski, *Virtual Machine Consolidation with Usage Prediction for Energy-Efficient Cloud Data Centers*. 2015 Ieee 8th International Conference on Cloud Computing, ed. C. Pu and A. Mohindra. 2015. 750-757.

65. Kim, N., J. Cho, and E. Seo, *Energy-credit scheduler: An energy-aware virtual machine scheduler for cloud systems.* Future Generation Computer Systems-the International Journal of Grid Computing and Escience, 2014. **32**: p. 128-137.

66. Fard, S.Y.Z., M.R. Ahmadi, and S. Adabi, *A dynamic VM consolidation technique for QoS and energy consumption in cloud environment.* Journal of Supercomputing, 2017. **73**(10): p. 4347-4368.

67. Dong, J.K., et al., *Energy-Saving Virtual Machine Placement in Cloud Data Centers*, in *Proceedings of the 2013 13th Ieee/Acm International Symposium on Cluster, Cloud and Grid Computing*, P. Balaji, D. Epema, and T. Fahringer, Editors. 2013. p. 618-624.

68. Cordeschi, N., et al., *Energy-efficient adaptive networked datacenters for the QoS support of real-time applications.* Journal of Supercomputing, 2015. **71**(2): p. 448-478.

69. Aryania, A., H.S. Aghdasi, and L.M. Khanli, *Energy-Aware Virtual Machine Consolidation Algorithm Based on Ant Colony System.* Journal of Grid Computing, 2018. **16**(3): p. 477-491.

70. Farahnakian, F., P. Liljeberg, and J. Plosila, *Energy-Efficient Virtual Machines Consolidation in Cloud Data Centers using Reinforcement Learning*, in *2014 22nd Euromicro International Conference on Parallel, Distributed, and Network-Based Processing*, M. Aldinucci, D. Dagostino, and P. Kilpatrick, Editors. 2014. p. 500-507.

71. Kord, N., H. Haghighi, and Ieee, *An Energy-Efficient Approach for Virtual Machine Placement in Cloud Based Data Centers*. 2013 5th Conference on Information and Knowledge Technology. 2013. 44-49.

72. Zarifzadeh, S., et al., *Joint range assignment and routing to conserve energy in wireless ad hoc networks.* Computer Networks, 2009. **53**(11): p. 1812-1829.

73. Mobius, C., W. Dargie, and A. Schill, *Power Consumption Estimation Models for Processors, Virtual Machines, and Servers.* Ieee Transactions on Parallel and Distributed Systems, 2014. **25**(6): p. 1600-1614.

74. Garg, S.K., et al., *Environment-conscious scheduling of HPC applications on distributed Cloud-oriented*

data centers. Journal of Parallel and Distributed Computing, 2011. **71**(6): p. 732-749.

75. Beloglazov, A., J. Abawajy, and R. Buyya, *Energy-aware resource allocation heuristics for efficient management of data centers for Cloud computing.* Future Generation Computer Systems-the International Journal of Escience, 2012. **28**(5): p. 755-768.

76. Moghaddam, F.A., P. Lago, and P. Grosso, *Energy-Efficient Networking Solutions in Cloud-Based Environments: A Systematic Literature Review.* Acm Computing Surveys, 2015. **47**(4).

77. Kong, F.X. and X. Liu, *A Survey on Green-Energy-Aware Power Management for Datacenters.* Acm Computing Surveys, 2015. **47**(2).

78. Arshad, R., et al., *Green IoT: An Investigation on Energy Saving Practices for 2020 and Beyond.* Ieee Access, 2017. **5**: p. 15667-15681.

79. Akhter, N. and M. Othman, *Energy aware resource allocation of cloud data center: review and open issues.* Cluster Computing-the Journal of Networks Software Tools and Applications, 2016. **19**(3): p. 1163-1182.

80. Lin, W.W., et al., *Experimental and quantitative analysis of server power model for cloud data centers.* Future Generation Computer Systems-the International Journal of Escience, 2018. **86**: p. 940-950.

81. Rong, H.G., et al., *Optimizing energy consumption for data centers.* Renewable & Sustainable Energy Reviews, 2016. **58**: p. 674-691.

82. Kaur, T. and I. Chana, *Energy Efficiency Techniques in Cloud Computing: A Survey and Taxonomy.* Acm Computing Surveys, 2015. **48**(2).

83. Jing, S.Y., et al., *State-of-the-art research study for green cloud computing.* Journal of Supercomputing, 2013. **65**(1): p. 445-468.

84. Cai, C., et al., *Energy-aware High Performance Computing - A Taxonomy Study*, in *2011 Ieee 17th International Conference on Parallel and Distributed Systems*. 2011. p. 953-958.

85. Bostoen, T., S. Mullender, and Y. Berbers, *Power-Reduction Techniques for Data-Center Storage Systems.* Acm Computing Surveys, 2013. **45**(3).

86. Dumitru, I., et al., *Increasing Energy Efficiency in Data Centers Using Energy Management*. 2011 IEEE/ACM International Conference on Green Computing and Communications. 2011. 159-65.

87. Yuan, X.L., et al., *Phase change cooling in data centers: A review.* Energy and Buildings, 2021. **236**.

88. Wu, C.M., R.S. Chang, and H.Y. Chan, *A green energy-efficient scheduling algorithm using the DVFS technique for cloud datacenters.* Future Generation Computer Systems-the International Journal of Escience, 2014. **37**: p. 141-147.

89. Mei, X., et al., *Energy-aware Task Scheduling with Deadline Constraint in DVFS-enabled Heterogeneous Clusters.* 2021.

90. Beloglazov, A., J. Abawajy, and R. Buyya, *Energy-aware resource allocation heuristics for efficient management of data centers for Cloud computing.* Future Generation Computer Systems, 2012. **28**(5): p. 755-768.

91. Goyal, T., A. Singh, and A. Agrawal, *Cloudsim: simulator for cloud computing infrastructure and modeling.* Procedia Engineering, 2012. **38**: p. 3566-3572.

92. Calheiros, R.N., et al., *CloudSim: a toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms.* 2011. **41**(1 %J Softw. Pract. Exper.): p. 23–50.