



Article Info

Received: 13th January 2023

Revised: 30th October 2023

Accepted: 25th November 2023

Department of Computer Science,
Faculty of Science, Federal University
Gusau, Gusau, Nigeria

*Corresponding author's email:

am2gus@yahoo.com

Cite this: *CaJoST*, 2023, 3, 273-279

Algorithm for allocating virtual computers in energy-friendly cloud data centers

Garba Muhammad* and Samaila Shuaibu

A huge proportion of centres that consume a lot of power, skyrocketing operating costs, and emit a disproportionate amount of carbon dioxide appeared due to the constantly increasing need for processing capacity. This necessitates finding a better approach to handling the energy crisis. This paper proposes a solution to address the energy crisis in cloud data centres, which consume a significant amount of power and emit a large amount of carbon dioxide. A power-conscious simulated engine allocation to physical machines reduces power consumption while maintaining an effective Class of Use. The proposed method is compared with other existing methods such as the Genetic Algorithm (GA), Modified Best Fit Decreasing Algorithm (MBFD), Random Selection (RS), and Minimum Migration Time (MMT) in terms of power usage. The proposed method outperforms the MBFD and GA procedures, showing an 8.55% and 46.81% improvement in power usage, respectively. The MBFD and GA procedures consume 389.5 and 812 kilowatts of energy, while the proposed method reduces this energy consumption. The proposed method is validated using the CloudSim simulator, which demonstrates its success as the fastest and most intelligent procedure compared to other methods. Additionally, the proposed method addresses the burdened host challenge quickly and provides valuable insights for future research by utilizing deep learning approaches.

Keywords: Cloudsim, Data Centres, Host, Simulation, Virtual Computers.

1. Introduction

The increasing power usage in the field of communication machinery, particularly in data centres, has indeed been a concern in recent years. Companies like Google, IBM, Microsoft, and others have expanded their data centre infrastructure to support the growing demand for cloud and grid processing capabilities. However, the operation of these data centres comes with significant power consumption and associated costs, as well as carbon dioxide emissions.

To address this issue, researchers have proposed various approaches to optimize power usage in data centres. One area of focus has been on identifying and eliminating redundant servers, which can consume a significant amount of power during idle or low-demand periods. Studies have shown that redundant servers can account for up to 70% of the power consumption during such periods. [1] [2] By accurately identifying these redundant servers and shutting them down when not needed, power consumption can be reduced. [3]

In the realm of cloud computing, methods for identifying redundant servers have been

developed with minimal overhead. It is important to avoid provisioning redundant servers for light workloads or no workloads, as they contribute to unnecessary power consumption. By implementing efficient server provisioning strategies and optimizing resource allocation through virtualization technologies, data centres can achieve significant reductions in power usage.

Controlling power usage in the cloud has become a popular research topic, not only to minimize energy consumption but also to adhere to green policies and meet the requirements and agreements set by consumers and facility providers. The adoption of virtualization technologies, along with careful management and relocation of virtualized resources, has proven effective in reducing power usage in data centres.

Overall, the focus on optimizing power usage in communication machinery, particularly in data centres, is driven by the need to reduce costs, minimize environmental impact, and align with sustainability goals. Through the implementation of efficient resource allocation, server

provisioning, and virtualization techniques, significant power savings can be achieved in these facilities.

[4] offered power aware procedure for distribution of VMs upon responses to three important enquiries:

- 1) What time is ideal for virtual computers transfer?
- 2) Which virtual computers will transfer?
- 3) Which location will the virtual computers transfer to?

In our work we opted for a matching technique, proposing best possible approach intended for the third enquiry in terms of Krill Procedure, a newly presented multi- investigative and known speedy procedure. For proper validation of the procedure, we employed the application of well-known less-laden host discovery processes, solitary-limit processes, excess-laden host discovery procedures, Interquartile Range algorithm (IQR) known either as the intermediate-dispersion or moderate half-half gauge of numerical distribution in factual markers [5].

Cloud computing has experienced rapid growth due to its ability to offer scalable computing resources and efficient processing and transmission of large amounts of data. Cloud processing involves the arrangement and consolidation of processing and transmission capabilities in an interactive manner, allowing for flexible and scalable handling of information infrastructure [6]. Cloud computing encompasses various types of services, including shared processing, parallel computing, and network processing. These technologies serve as a foundation for cloud processing and enable the provisioning of hardware resources, basic software services, programming tools, application services, and storage space over the internet [7]. One key aspect of cloud computing is that it is typically provided by a third-party service provider who owns and manages the underlying infrastructure. This means that users can focus on utilizing the services provided by the cloud without having to worry about the maintenance and management of the physical infrastructure.

Indeed, the definition of cloud computing has been standardized based on the input and research of various experts. The cloud computing standardization white paper from China provides a common understanding of cloud computing. According to this white paper, cloud computing is defined as a model that offers on-demand access to shared, flexible, and dynamically scalable real

and virtual computing resources. Users can request these resources through self-service portals and access them via web interfaces.

The white paper highlights several key elements that constitute the cloud computing model. These elements include:

1. Responsibilities and events: This refers to the roles and responsibilities of different stakeholders involved in cloud computing, such as cloud service providers, users, and administrators. It also covers the events and interactions that occur within the cloud computing environment.
2. Cloud capacity categories and cloud usage division: Cloud computing resources can be categorized based on their capacity and capabilities, such as processing power, storage capacity, and network bandwidth. The white paper addresses how these resources are divided and allocated to different users or applications.
3. Cloud utilization patterns: Cloud computing supports various utilization patterns, such as Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS). These patterns define the level of abstraction and services provided by the cloud, ranging from infrastructure-level resources to fully managed applications.
4. Cloud computing concerns: The white paper acknowledges that there are common concerns and considerations related to cloud computing. These may include security, privacy, data sovereignty, interoperability, and compliance with regulations. Addressing these concerns is essential to ensure the trust and confidence of cloud computing users [8].

The service model upheld that cloud processing exist as Software as a Service (SaaS) [9], Platform as a Service (PaaS) [10], and Infrastructure as a Service (IaaS) [11]. Placement approach characterized processing into: personal cloud, neighbourhood cloud, open cloud and mix cloud [12].

[13] From the foregoing, it is evident to claim that Cloud processing in terms of uses bear some resemblance with grid processing in some ways and clearly distinct in other ways.

In the present, Cloud suppliers often known as data facilities, engaged in offering many vigorous and important supplies in the direction of the clients. The data facility support clients within and beyond to use the facilities they provided directly through implicit requests. Notable Information

Technology giants, such as Facebook, Amazon, Google, Microsoft, Yahoo, and Oracle, have their own data facilities to provide cloud services. These companies leverage their infrastructure and expertise to offer a wide range of cloud computing resources and services to customers [14]. In the context of cloud computing, there are indeed various challenges that have been identified and discussed. Some of these challenges include: cost, service provider reliability and downtime, data privacy, energy consumption and sustainability, virtual machine (VM) migration, server consolidation and energy efficiency, vendor lock-in [13].

In the present, optimizing energy efficiency in cloud data centres has become a priority for IT units. However, there are still significant environmental concerns associated with cloud computing, including greenhouse gas emissions, carbon dioxide emissions, and the generation of electronic waste (e-waste). Data centre machines are known to be the most energy-intensive components in cloud infrastructure. When a machine operates at its peak, it generates a significant amount of heat. To manage this heat and prevent overheating, the load of heavily loaded machines can be distributed to either redundant machines or machines with lower workloads. Achieving energy efficiency in cloud computing requires a combination of strategies that adhere to regulatory policies and environmental principles while reducing energy consumption [15].

[21]–[23] submitted that energy efficiency-Performance Trade-off has significance in cloud processing concept, having energy organization controls focusing on prospective effect on energy productivity and execution is essential. Quantities Turning redundant devices to small energy positions provide remarkable drop of hardware rate or alertness; server merging causes inner conflict; temperature-friendly job distribution acquires substantial costs. Of significance is the server's power-effectiveness due to its effects on the electrical and the heat reduction expenses responsible for the main unit of the overall expenses incurred in running a data hub. The ever-rising power expenses shoot up the overall expenses of Proprietorship and minimize the Profit on Capital Spending of cloud structures. Subsequently, cloud suppliers faced the difficulty of reducing the running data hub expenses at the same time as attaining a great deal amid energy reserves and product execution

Joining a cloud service may be relatively straightforward, but transitioning to a different provider or exiting the cloud environment can be

more complex and costly. This situation is known as vendor lock-in. Vendor lock-in can occur when a client becomes heavily dependent on a specific cloud service provider's technologies, APIs, or proprietary solutions. This dependence can make it challenging to switch to another provider without incurring additional expenses or encountering compatibility issues. It can also raise concerns about the portability of data and applications [24]–[26]. Interoperability and portability are related concepts that aim to address the challenges of vendor lock-in. Interoperability refers to the ability of different systems or platforms to exchange and use information effectively. In the context of cloud computing, interoperability involves ensuring that clients can access and utilize their data, applications, and control tools across multiple cloud service providers and infrastructures. One of the problems related to interoperability and portability is the lack of common standards and compatibility of interfaces and APIs across different cloud service providers. Inconsistent or proprietary APIs can hinder the seamless integration and portability of applications and data between different cloud environments.

Efforts are being made by industry organizations, standards bodies, and cloud service providers themselves to address these challenges. Developing and adhering to common standards, open APIs, and data formats can promote interoperability and portability in the cloud computing ecosystem. These initiatives aim to provide clients with more flexibility and freedom of choice when it comes to selecting and switching between cloud service providers [24]–[27].

2. Methodology

2.1 Energy consumption

Watts is the unit for calculating power, which is the proportion of finished job. Power could be the proportion of energy utilization. Therefore, energy is the overall job executed during a period.

$$p = \frac{E}{T} \quad (1)$$

where, p is power, T is time and E is energy consumed in time T

Energy utilization is calculated through different techniques. The most common technique calculates energy utilization from panel, though simple, does not allow for examining or indicating the point of the real energy utilization. Similarly, energy consumption is calculated from simulations devoid of applying some hardware elements [28].

2.2. Materials and Methods

Cloud computing platform possess two kinds of virtual machines, the mixed and energetic virtual machines. In this work Krill designate simulated machines and the food attraction designate physical hosts. Equation (2) represents the collection of simulated machines in the S set.

$$V_{C_r} = \{v_1, v_2, v_3, \dots, v_u\} \quad (2)$$

u designates the overall figure of simulated machines and equation (3) represents the collection of physical hosts

$$H_{C_r} = \{h_1, h_2, h_3, \dots, h_m\} \quad (3)$$

m designates the overall figure of physical hosts

The work fixes the challenge of assigning simulated machines to corporal hosts tactically that every simulated machine is mapped to exactly a corporal host and maintaining the least corporal hosts to be operating. In dealing with the challenge of assigning simulated machines to corporal hosts, every solution was represented in a rectangular collection of simulated machine reference number and corresponding corporal host to which simulated machine reference number is assigned as illustrated in the form of an array represented by [1].

Table1: Illustration of an answer

Vc no	0	1	2	...	U
Crp_m no	7	3	5	...	M

In the table1, simulated machine with reference number 0 is assigned to corporal machine number 7. We have in this technique associated a rectangular collection to every krill to denote the distance traversed by the krill. Reference number 0 is the initial position where any krill begins its movement while the terminal reference number is the final position of food source.

The following is the algorithm of the technique:

Step 1: Initialize algorithm's parameters

Step 2: Create the initial population of krill

Step 3: Fitness evaluation: Each krill individual is evaluated according to its position

Step 4: Motion Calculation:

- 4.1 Motion induced by the presence of other individuals
- 4.2 Foraging motion
- 4.3 Physical diffusion
 - 4.3.1 IQR detects host overloaded
 - 4.3.2 MMT or RS select the VMs to migrate

Step 5: Update the krill position in the search space

Step 6: Go to step 3 until the best solution is obtained

Step 7: Stop

2.3. Choice machinery

The rth krill chooses physical machine according to the sequence of corporal machines described in the array, assigning simulated engine to the supplier j base on equation (3), $d_q(t)$ is the remoteness to the nutrition source q at time t.

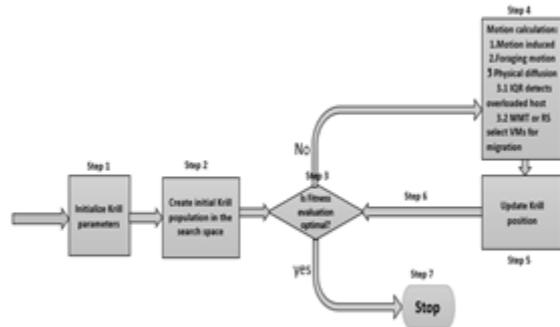


Figure 1. Krill Herd algorithm

$$p_r^\theta(d) = \frac{[d_r(t)]^{\alpha} * [u_r]^{\beta}}{\sum [d_r(t)]^{\alpha} * [u_r]^{\beta}} \quad (4)$$

Table2: Parameters applied in the algorithm

Parameter	Description
Vc	Virtual machine
Pm	Physical machine
v_rCPU	Needed processor for virtual machine (r)
V_rmem	Needed memory for virtual machine (r)
P_qcpu	Needed processing capacity for physical machine (q)
p_qmem	Needed memory for physical machine (q)
p_qccpu	Cumulative workload of physical machine (q)
P_qcmem	Cumulative memory utilized for physical machine (q)

Where u_q is the innate volume of the source q, α is the parameter managing the impact of $d_r(t)$, and β is a parameter to manage the impact of u_r . A global update occurs at the time each krill got an answer to the VM distribution challenge, krill foraging remoteness related to entire physical machines chosen by the finest krill possessing least power usage may be revised by equation (5):

$$d_r^{new} = (1 - \partial)d_r + \Delta d_r \quad (5)$$

From the above relation, Δd_r rises, and ∂ designates the factor of distance towards the food

and the value of the search distance is controlled to stimulate solution.

A suitability function was applied to determine if a suitable answer can be obtained from the solution in terms of the defined parameters responsible for excellent solution. Using the function, value of every solution is evaluated to pick the finest answer, extreme or least value according to the limit assignment strategy is regarded as the greatest suitable solution. Consequently, the suitability function is defined by equation (5) below:

$$suitability = \frac{1}{\sum_1^u E(p_r)} \quad (6)$$

2.4. Experimentation

We employ the CloudSim tools to develop and simulate cloud computing structures and to create applications. Data centres, virtual machines, resource policies, and application delivery strategies are few of the cloud computing system constituents that may be moulded using CloudSim, in addition to their behaviour as a whole. [4]. Study of the work presented in [29] showed the bio-inspired algorithm proposed fits this area well. Hence, a more precise assessment of the method was replicated, and the result related to that of the mentioned algorithm. Consequently, for proper assessment of the algorithm, data from available research are used. In our simulation we used IQR to be laden host discovery, arbitrary policy quest, and random selection, and minimal migration time algorithms for virtual computer choice during the multiple rounds. Table 6 showed the progress in energy usage by the proposed algorithm correlated to the other methods. The results are showed in the figures; 3 and 4 portraying the energy consumption of the three algorithms.

Table 3: Data centre specification

Experiment	VMs	Physical Host	Data centre
	1175	800	1

Table 4: Virtual machine specification

Experiment	RAM(MB)	CPU	MIPS
	613-3840	1	500-2500

Table 5: Physical machine specification

Experiment	RAM(MB)	CPU	MIPS
	4096	2	1860-2660

Table 6: Workload specification

Experiment	MIPS(Tasks)	Number of Tasks
	500-2500	1175

3. Results and Discussion

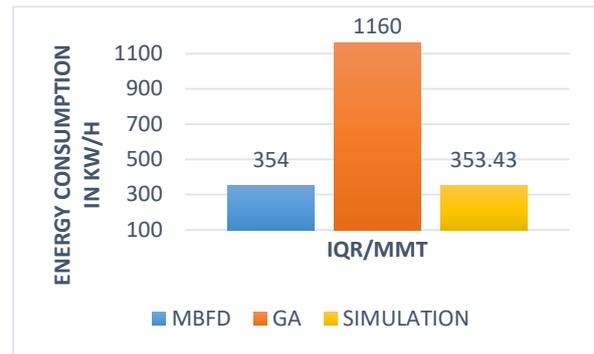


Figure 2. Least energy consumption of the algorithms for simulated machine addition into corporal hosts using IQR/MMT.

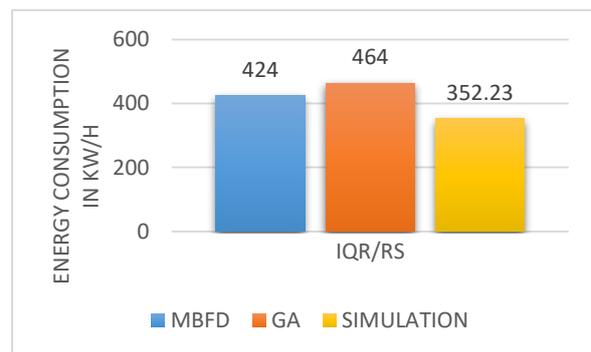


Figure 3. Least energy consumption of the algorithms for simulated machine addition into corporal hosts using IQR/RS.

Table 7: Average progress of energy usage in the proposed method correlated to GA method

	MMT	RS
IQR	69.53	24.09
Average	46.81%	

Table 8: Average progress of energy usage in the proposed method correlated to MBFD method

	MMT	RS
IQR	0.16	16.93
Average	8.55%	

Table 9: Smallest energy usage in the proposed method in every algorithm for adding virtual computer into physical hosts

Energy consumption	Algorithm		
352.83	MMT-RS	IQR-RS	PROPOSED
812	MMT-RS	IQR-RS	GA
389.5	MMT-RS	IQR-RS	MBFD

Table10: Average progress of proposed method correlated to the other methods

MBFD	GA
8.55%	46.81%

On the results, tables 7, 8 and 9 indicate energy usage. Energy usage is one of the yardsticks for evaluating virtual computer placement into the corporal host normally calculated in Kilo watt per hour. The figures have portrayed that the proposed method did well to minimize energy usage at 8.55% and 46.81% correlated to MBFD and GA methods. Additionally, table 7 depicts energy usage using IQR as overburdened host discovery strategy, the same is the case in table 8. Random choice and least relocation period were both used for virtual computer choice throughout the multiple execution of the experiments. Tables 9 and 10 pictured the smallest energy usage in every method for placement of virtual computers to the corporal hosts.

4. Conclusion

Power utilization remains a notable bottleneck in today's cloud processing putting cloud reserve Provisionals under pressure. Recent study suggests that reliable solution could be obtainable. A technique was proposed that aimed at controlling energy usage through optimizing the number of virtual computers and switching off static waiters. The results indicate that proper combination of, and the choice of ideal virtual computer relocation policies can guaranty power efficacy. Correlating the proposed method to the other methods with IQR as laden host discovery. Chance selection, and least relocation spell strategies as virtual computer choice for placement into the hosts the multiple circles it can be concluded that energy usage has improved by 8.55% and 46.81%, individually. Furthermore, applying profound learning procedures may promptly address the laden host challenge and may be an insight for further study.

Conflict of interest

We declare that there is no conflict of interest.

References

- [1] M. Soltanshahi, R. Asemi, and N. Shafiei, "Energy-aware virtual machines allocation by krill herd algorithm in cloud data centers," *Heliyon*, vol. 5, no. 7, Jul. 2019. doi.org/10.1016/j.heliyon.2019.e02066
- [2] D. Liu, X. Sui, and L. Li, "An energy-efficient virtual machine placement algorithm in cloud data center," *2016 12th Int. Conf. Nat. Comput. Fuzzy Syst. Knowl.*

- [3] A. Beloglazov, "Energy-Efficient Management of Virtual Machines in Data Centers for Cloud Computing," *PhD Thesis, Univ. Melb.*, no. February, pp. 1--232, 2013. doi.org/10.1016/j.future.2011.04.017
- [4] A. Beloglazov, J. Abawajy, and R. Buyya, "Energy-aware resource allocation heuristics for efficient management of data centers for cloud computing," *Futur. Gener. Comput. Syst.*, vol. 28, no. 5, pp. 755–768, 2012. doi.org/10.1016/j.future.2011.04.017
- [5] B. Guenter, N. Jain, and C. Williams, "Managing cost, performance, and reliability tradeoffs for energy-aware server provisioning," *Proc. - IEEE INFOCOM*, pp. 1332–1340, 2011. doi:10.1109/INFOCOM.2011.5934917
- [6] L. Qian *et al.*, "Overview of Cloud Computing," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 677, no. 4, 2019. Lin Qian *et al* 2019 IOP Conf. Ser.: Mater. Sci. Eng. 677 042098
- [7] B. Heydari and M. Aajami, "Providing a new model for discovering cloud services based on ontology," *Eng. Technol. Appl. Sci. Res.*, vol. 7, no. 6, pp. 2268–2272, 2017. doi.org/10.48084/etasr.1577
- [8] B. Holzman *et al.*, "Hepcloud, a new paradigm for hep facilities: Cms amazon web services investigation," *Comput. Softw. Big Sci.*, vol. 1, no. 1, p. 1, 2017. doi:10.1007/s41781-017-0001-9
- [9] A.-S. K. Pathan and M. M. Z. E. Mohammed, "Building Customer trust in cloud computing with an ICT-enabled global regulatory body," *Wirel. Pers. Commun.*, vol. 85, no. 1, pp. 77–99, 2015. doi:10.1007/s11277-015-2729-z
- [10] J. Baek and C. Kim, "An energy-efficient video transport protocol for personal cloud-based computing," *J. Real-Time Image Process.*, vol. 12, no. 2, pp. 303–310, 2016. doi:10.1007/s11554-014-0475-1
- [11] F. Cannas, "The VAT treatment of cloud computing: legal issues and practical difficulties," *World J. VAT/GST Law*, vol. 5, no. 2, pp. 92–110, 2016. doi:10.1080/20488432.2016.1236622
- [12] J. B. Abdo and J. Demerjian, "Evaluation of mobile cloud architectures," *Pervasive Mob. Comput.*, vol. 39, pp. 284–303, 2017. doi: 10.1016/j.pmcj.2016.12.003
- [13] T. Kaur, "COMPREHENDING CLOUD COMPUTING FOR REALIZING ENERGY," vol. 21, no. 6, pp. 532–541, 2019. ISSN: 0374-8588 Volume 21 Issue 6,

- October 2019
- [14] F. Teng, "Management of Data and Scheduling of Tasks on Architecture Distributees," PhD Thesis, École Centrale: A University Institution, Paris, 2011.
- [15] L. Lefèvre and A.-C. Orgerie, "Designing and evaluating an energy efficient cloud," *J. Supercomput.*, vol. 51, no. 3, pp. 352–373, 2010. DOI 10.1007/s11227-010-0414-2
- [16] N. J. Kansal and I. Chana, "Existing load balancing techniques in cloud computing: a systematic review," *J. Inf. Syst. Commun.*, vol. 3, no. 1, p. 87, 2012.
- [17] R. Wang, W. Le, and X. Zhang, "Design and implementation of an efficient load-balancing method for virtual machine cluster based on cloud service," 2011. doi:10.1049/cp.2011.1015
- [18] F. Alharbi, Y. C. Tian, M. Tang, W. Z. Zhang, C. Peng, and M. Fei, "An Ant Colony System for energy-efficient dynamic Virtual Machine Placement in data centers," *Expert Syst. Appl.*, vol. 120, pp. 228–238, 2019. doi: 10.1016/j.eswa.2018.11.029
- [19] Y. C. Lee and A. Y. Zomaya, "Energy efficient utilization of resources in cloud computing systems," *J. Supercomput.*, vol. 60, no. 2, pp. 268–280, 2012. doi:10.1007/s11227-010-0421-3
- [20] S. Srikanthiah, A. Kansal, and F. Zhao, "Energy aware consolidation for cloud computing," 2008. DOI: 10.1002/spe.2585
- [21] A. Beloglazov, J. Abawajy, and R. Buyya, "Energy-aware resource allocation heuristics for efficient management of data centers for Cloud computing," *Futur. Gener. Comput. Syst.*, vol. 28, no. 5, pp. 755–768, 2012. doi: 10.1016/j.future.2011.04.017
- [22] D. Dharwar, S. S. Bhat, V. Srinivasan, D. Sarma, and P. K. Banerjee, "Approaches towards energy-efficiency in the cloud for emerging markets," in *2012 IEEE International Conference on Cloud Computing in Emerging Markets (CCEM)*, 2012, pp. 1–6. doi:10.1109/ccem.2012.6354599
- [23] H. Yuan, C.-C. J. Kuo, and I. Ahmad, "Energy efficiency in data centers and cloud-based multimedia services: An overview and future directions," in *International Conference on Green Computing*, 2010, pp. 375–382. doi:10.1109/greencomp.2010.5598292
- [24] R. Buyya, A. Beloglazov, and J. Abawajy, "Energy-efficient management of data center resources for cloud computing: a vision, architectural elements, and open challenges," *arXiv Prepr. arXiv1006.0308*, 2010. PhD Thesis, Feb. 2013, Department of Computing and Information Systems, THE UNIVERSITY OF MELBOURNE
- [25] M. Hussin, Y. C. Lee, and A. Y. Zomaya, "Priority-based scheduling for large-scale distribute systems with energy awareness," in *2011 IEEE Ninth International Conference on Dependable, Autonomic and Secure Computing*, 2011, pp. 503–509. doi:10.1109/DASC.2011.96
- [26] L. Luo, W. Wu, D. Di, F. Zhang, Y. Yan, and Y. Mao, "A resource scheduling algorithm of cloud computing based on energy efficient optimization methods," in *2012 International Green Computing Conference (IGCC)*, 2012, pp. 1–6. doi:10.1109/IGCC.2012.6322251
- [27] Q. Zhang, L. Cheng, and R. Boutaba, "Cloud computing: state-of-the-art and research challenges," *J. internet Serv. Appl.*, vol. 1, no. 1, pp. 7–18, 2010. doi:10.1109/IGCC.2012.6322251
- [28] G. Anithakrishna and M. Mohankumar, "An Empirical Study on Energy Estimation Tools," vol. 21, no. 7, pp. 104–109, 2019. doi:10.1109/I GCC.2012.6322251