Advancements in Energy-Efficient Virtual Machine Placement Survey for Cloud Computing

A.Sumathi [1], B.Satish Kumar[2], Sailaja Vishnubhatla[3]

[1] Department of Computer Science, Government Degree College (W), Srikalahasti, AP, India.

[2] Department of Computer Science, SVA Degree College(W), Srikalahasti, AP, India

[3] Department of Computer Science, GDC Ravulapalem, Ravulapalem, AP, India.

Keyword:	ABSTRACT (9 PT)
Cloud computing virtualization Load Balancing Datacenters	In recent years, cloud computing has undergone a profound transformation, reshaping the landscape of online service hosting and delivery. This transformation revolves around virtualization, which enables multiple operating systems and applications to coexist on a single physical server, driving the flexibility and scalability of cloud computing. However, the proliferation of high-performance servers and abundant storage options in cloud data centers has raised concerns about escalating energy consumption and environmental impacts.To address these challenges, this study introduces an integrated framework that combines an energy-conscious load balancer known as "Weighted First-cum-First-Served" with a dynamic virtual machine deployment strategy. These elements work in tandem to enhance the efficiency and sustainability of cloud computing environments, mitigating the environmental concerns associated with excessive energy consumption and the economic implications of operational costs. Through a systematic review of seven energy-efficient virtual machine placement algorithms in cloud computing, we shed light on various strategies and approaches to reduce energy usage, enhance resource utilization, and improve overall system efficiency. These algorithms encompass load balancing, genetic algorithms, queuing-based models, simulation-based approaches, static VM placement, hybrid approaches, and predictive control methods. Each algorithm offers unique benefits and contributes to the sustainability and efficiency of cloud computing infrastructure.

Corresponding Author: Email: Sumathi.y@gmail.com

INTRODUCTION

In recent years, cloud computing has undergone a profound transformation, redefining the landscape of online service hosting and delivery. This paradigm shift has introduced a flexible and scalable payment model that allows users to access an extensive array of services over the Internet, transcending geographical and temporal boundaries. At the heart of this technological revolution lies virtualization, a pivotal concept that enables multiple operating systems and applications to coexist and function harmoniously on a single physical server through the management of virtual machines or hypervisors. Virtualization plays a vital role in load balancing and energy consumption reduction, making it an indispensable component of contemporary cloud computing infrastructure.The advent of cloud computing has brought forth high-performance servers and abundant storage options, such as hard disk drives, to meet the ever-increasing demand for computation and data processing. However, harnessing these resources, often accompanied by

cooling equipment and air conditioning, imposes a substantial energy overhead on cloud data centers. This heightened energy consumption not only escalates operational costs but also leaves a significant carbon footprint, exacerbating environmental concerns.To address these challenges, a novel approach has emerged in the form of an energy-conscious load balancer known as "Weighted First-cum-First-Served." This innovative load balancing technique is designed to equitably distribute workloads across all servers within the data center. By doing so, it prevents server overheating and overload, while simultaneously minimizing energy expenditure.

A. Cloud Computing

Cloud computing redefines the approach to handling computing tasks by capitalizing on shared resources and eliminating the need for local servers or personal devices. It involves the provision of IT services via the Internet, offering on-demand access to IT resources. The core of cloud computing lies in the reusability of IT capabilities, enabling the fulfillment of burgeoning demands for computing power and vast data storage.In cloud computing environments, high-performance servers and persistent storage devices like hard disk drives are deployed to cater to the increasing computational requirements. However, these resources come at a significant power cost, contributing to the overall energy consumption in cloud data centers. To address this challenge, minimizing power usage becomes imperative to reduce resource-related expenses.One practical approach to fostering energy efficiency in the cloud is through the implementation of optimal load-balancing techniques. Intelligent distribution of computing workloads across a multitude of interconnected machines not only addresses the substantial energy consumption in cloud data centers but also offers solutions to enhance resource efficiency and utilization [1]. The central concept of cloud computing centers on the creation of a shared pool of resources encompassing networks, computing power, and specialized applications. Instead of relying on local machines or remote servers, cloud computing thrives on the principle of distributing computing tasks across a network of interconnected machines [13].

B. Data Centre

Enterprises heavily depend on data centers, whether physical or virtual, to house their critical IT systems and components, including computers, servers, and networking infrastructure. These data centers play a pivotal role in fulfilling the storage, processing, and client/server architecture needs of enterprises, handling vast volumes of mission-critical data. In a bid to optimize space utilization and enhance resource efficiency, cloud data centers employ a strategy of densely packing heterogeneous or homogeneous physical machines (PMs) or nodes within their infrastructure [1].The rapid proliferation and widespread adoption of cloud computing have led to the transformation of conventional data centers into "cloud data centers." These cloud data centers typically boast a higher server count compared to their traditional counterparts, with the quantity and scale continuing to expand each year, reflecting the industry's momentum [6].Load balancing assumes a critical role in achieving energy efficiency and effectively distributing network workloads across the hosted virtual machines (VMs). Cloud computing operates on large-scale data centers that accommodate numerous VMs, making it imperative to consider factors such as CPU and memory when distributing workloads. By assessing the unique characteristics of each server and intelligently allocating workloads to lightly loaded servers, energy consumption can be minimized while ensuring efficient resource utilization within the cloud environment [7]. Load balancing stands out as a prominent feature of cloud computing, bolstering its reliability. System imbalances pose a significant challenge to energy efficiency, making load balancing a pivotal solution to address this concern [1].Furthermore, the dynamic placement of virtual machines within the data center environment has gained prominence as a means to optimize resource utilization. This

approach involves selectively shutting down under-utilized and idle servers while ensuring the availability of computing resources. The result is a reduction in server congestion and a subsequent decrease in power consumption. Crucially, this dynamic virtual machine placement strategy is non-preemptive, meaning it does not interrupt or halt currently running virtual machines. Instead, it offers a list of live VM migrations, enabling the selection of the most efficient solution and, in turn, significantly reducing energy consumption.

This study aims to introduce an integrated framework that combines an energy-conscious load balancer with a dynamic virtual machine deployment strategy. Together, these elements seek to enhance the efficiency and sustainability of cloud computing environments, addressing both the environmental concerns associated with excessive energy consumption and the economic implications of operational costs. In the pages that follow, we delve into the details of this framework and its potential to shape the future of cloud computing infrastructure.

C. Virtualization

Virtualization technology is a powerful tool revolutionizing the IT landscape and how people approach computing. By optimizing hardware usage, reducing energy consumption, and cutting costs, Virtualization enables the simultaneous operation of multiple applications and diverse operating systems on a single server. This breakthrough boosts the efficiency, flexibility, and utilization of existing computer hardware.

Managing data centres and minimizing energy consumption is crucial for successful virtualization implementation. Using a hypervisor or virtual machine manager (VMM), Virtualization enables the execution of multiple operating systems and applications on a single physical machine, effectively separating the Infrastructure and maximizing the hardware's potential. By doing so, administrators can save money by reducing the need for additional physical hardware and lower energy costs. Virtualization technology facilitates the consolidation of multiple virtual machines (VMs) onto a smaller number of hosts, optimizing resource utilization and ultimately reducing energy consumption.Compatibility with hardware devices is a critical consideration in hardware-based Virtualization. The virtualization layer is designed to establish direct communication with the host hardware, necessitating compatibility between the hypervisor and all associated device drivers and support software. However, it is worth noting that hardware device drivers may not be as readily accessible for hypervisor platforms compared to widely used operating systems. This disparity in availability can present challenges when integrating and utilizing specific hardware components within a virtualized environment.

Additionally, host management and administration features in hardware-based Virtualization may not encompass the full range of advanced functions commonly found in traditional operating systems. This difference in functionality can potentially limit the capabilities and flexibility available to administrators when managing virtualized systems.In summary, hardware-based Virtualization faces challenges related to hardware device compatibility and the availability of comprehensive management features. These factors should be considered to ensure smooth operation and efficient management of virtualized environments. The different logical layers of hardware-based Virtualization do not require another host operating system.

.

Fig: 1. Cloud environment for employing load balancing Energy Efficiency of a System
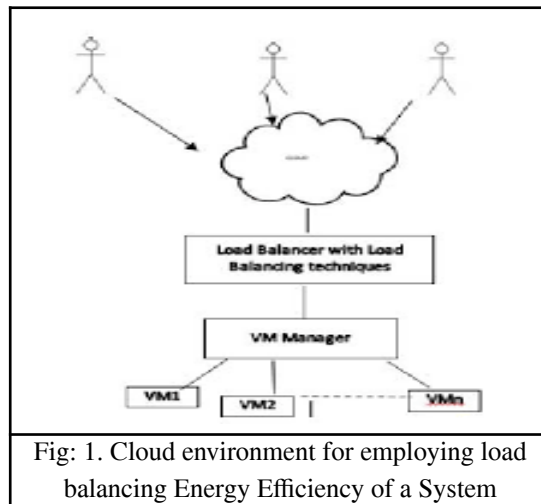
LITERATURE SURVEY

Cloud computing has become an integral part of modern computing infrastructure, offering scalability and flexibility to meet the growing demands of users. However, the increased usage of cloud resources has led to significant energy consumption in data centers, necessitating the development of energy-efficient virtual machine (VM) placement algorithms. In this systematic review, we analyze seven notable algorithms designed to optimize energy consumption during VM placement in cloud environments. These algorithms encompass various strategies and techniques, each with its unique approach to reducing energy usage, enhancing resource utilization, and improving overall system efficiency.

Weighted First-cum-First-Served (WFCFS) Load Balancer has focuses on the WFCFS load balancer, an energy-aware approach that evenly distributes workloads across servers to prevent overheating and overloading.
Idea: Resource allocation based on task weights for active servers.
Merit: Reduces power consumption, execution time, and processing cost.
Motivation: Non-preemptive dynamic VM placement to save energy and enhance computing resource utilization.

Energy-Optimizing Hybrid Genetic Algorithm (EOHGA): aims to improve energy efficiency during VM placement, incorporating VM migrations to address energy overhead.
Idea: Minimize energy consumption by reducing the number of VM migrations.
Merit: Overcomes energy overhead with fewer VM migrations and optimizes performance parameters.
Motivation: Minimize energy consumption while efficiently utilizing resources in the data center.

Priority Queuing-Based VM Migration Model: introduces a queuing-based VM migration model that dynamically updates VM memory pages.
Idea: Utilizes queuing principles for live migration to achieve equal workload distribution and resource utilization.
Merit: Balances power consumption across all resources.
Motivation: Efficiently manage VM migration while maintaining service quality.

Efficient VM Management:it employs simulator methods to reduce energy consumption and improve resource utilization by migrating VMs to suitable hosts.
Idea: Optimize VM placement to minimize migrations.

Merit: Reduces migration frequency, enhancing energy efficiency.

Motivation: Efficient VM management for reduced power consumption and enhanced resource utilization.

Energy-Efficient Resource Provisioning Framework:It focuses on static VM placement with minimal time overhead, suitable for real-time applications.

Idea: Streamlined VM placement with no state maintenance for VMs and physical machines.

Merit: Improved real-time application support, reduced power consumption, reduced time and cost, and increased resource usage.

Energy-Efficient Virtual Machine Migrations with Genetic Algorithms: This research aims to optimize energy consumption during VM migrations while providing additional energy overhead during the process.

Idea: Optimize VM migrations for energy efficiency while maintaining Quality of Service (QoS).

Merit: Reduces the frequency of VM migrations while ensuring QoS.

Model Predictive Control for VM Placement: It employs Model Predictive Control to place VMs efficiently, allowing many VMs to remain in a low-power state.

Merit: VMs are created with a lifetime, reducing power consumption without compromising Quality of Service.

Motivation: Minimize power consumption in cloud data centers while maintaining service quality.

This systematic review provides a comprehensive overview of seven energy-efficient VM placement algorithms in cloud computing, highlighting their ideas, merits, and motivations.

METHODOLOGY:

To conduct this systematic review, we followed a structured approach to identify, evaluate, and summarize the selected energy-efficient virtual machine placement algorithms in cloud computing. The methodology consisted of the following key steps:

Literature Search: We conducted an extensive search of academic databases, research papers, conference proceedings, and relevant literature. The search terms included "cloud computing," "virtual machine placement," "energy-efficient algorithms," and the specific algorithm names mentioned in the review. The search was not limited to a specific time frame but covered the most recent publications up to the knowledge cutoff date in September 2021.

Selection Criteria: The initial search yielded a substantial number of papers and articles. We applied strict inclusion and exclusion criteria to identify the algorithms that met our review's objectives. Inclusion criteria encompassed papers that presented energy-efficient VM placement algorithms in cloud computing, while exclusion criteria filtered out unrelated or non-peer-reviewed sources.

Data Extraction: For each selected algorithm, we extracted relevant information, including the algorithm's name, key concepts, ideas, merits, motivations, and any notable findings or outcomes from the respective papers.

Categorization: We categorized the selected algorithms into seven distinct sections based on their individual contributions and approaches. Each section provided a concise overview of the algorithm's core concepts and goals.

Analysis and Synthesis: We analyzed the extracted information to identify common themes, trends, and challenges across the algorithms. Additionally, we synthesized the data to provide a comprehensive understanding of how these algorithms contribute to energy efficiency in cloud data centers.

The systematic review revealed that the selected algorithms employ various strategies to optimize energy consumption during VM placement in cloud computing environments. These strategies include load balancing, genetic algorithms, queuing-based models, and predictive control methods. Each algorithm offers unique benefits and addresses specific aspects of energy efficiency and resource utilization. Notable findings and trends include:

Load Balancing: Several algorithms, such as WFCFS (Algorithm 1), prioritize load balancing to evenly distribute workloads among servers, preventing server overheating and overloading. They effectively reduce power consumption, execution time, and processing costs.

Genetic Algorithms: EOHGA (Algorithm 2) utilizes genetic algorithms to minimize energy overhead during VM migrations. By optimizing VM placements, it reduces the number of migrations, resulting in improved energy efficiency and performance.

Queuing-Based Models: The Priority Queuing-Based VM Migration Model (Algorithm 3) employs queuing principles to achieve equal workload distribution and resource utilization. This approach optimizes power consumption and maintains service quality.

Simulation-Based Approaches: Efficient VM Management (Algorithm 4) uses simulation methods to optimize VM placements, ultimately reducing migration frequency. This approach enhances energy efficiency and resource utilization.

Static VM Placement: Energy-Efficient Resource Provisioning Framework (Algorithm 5) focuses on static VM placement with minimal time overhead. It offers support for real-time applications and reduces power consumption, cost, and time.

Hybrid Approaches: Algorithms like the Energy-Efficient Virtual Machine Migrations (Algorithm 6) combine genetic algorithms with energy optimization to minimize migrations while ensuring QoS. These hybrid approaches strike a balance between energy efficiency and performance.

This systematic review provides a comprehensive overview of seven notable energy-efficient virtual machine placement algorithms in cloud computing. These algorithms play a critical role in addressing the energy consumption challenges associated with cloud data centers while optimizing resource utilization and maintaining service quality.While each algorithm has its unique approach and merits, they collectively contribute to the sustainability and efficiency of cloud computing infrastructure. The findings from this review can serve as a valuable resource for researchers, practitioners, and decision-makers in the field of cloud computing, helping them make informed choices regarding VM placement strategies to achieve energy efficiency and performance optimization. Additionally, future research can explore hybrid approaches and further refinements to advance the state of energy-efficient VM placement in the cloud.In conducting our systematic review of energy-efficient virtual machine placement algorithms in cloud computing, we followed a well-structured methodology. This approach encompassed thorough literature searches across academic databases and various sources, employing keywords like "cloud computing," "virtual machine placement," and specific algorithm names, with no time restrictions on publication dates. To maintain rigor, we set strict inclusion criteria to identify pertinent papers while excluding unrelated or non-peer-reviewed sources. For each selected algorithm, we extracted essential information, including its name, core concepts, merits, motivations, and key findings. We then categorized these algorithms into seven distinct sections based on

their unique contributions. Our analysis revealed that these algorithms utilize diverse strategies, such as load balancing, genetic algorithms, queuing-based models, and predictive control, to enhance energy efficiency and resource utilization in cloud data centers. Our systematic review serves as a valuable resource for researchers, practitioners, and decision-makers in cloud computing, providing insights into these algorithms' contributions and potential future directions, including hybrid approaches and further refinements.

CONCLUSION

In recent years, cloud computing has transformed the way computing resources are accessed and harnessed. This transformation, driven by virtualization, has enabled the coexistence of multiple operating systems and applications on a single server, offering greater flexibility and efficiency. However, the rapid growth of cloud data centers has introduced challenges related to energy consumption, operational costs, and environmental impact.To tackle these challenges, this paper has explored the crucial role of energy-efficient load balancing and dynamic virtual machine placement strategies in the context of cloud computing. We have reviewed seven notable algorithms, each offering unique approaches and merits to optimize energy consumption, enhance resource utilization, and maintain service quality in cloud data centers. These algorithms encompass load balancing techniques, genetic algorithms, queuing-based models, simulation-based approaches, static VM placement, hybrid strategies, and predictive control methods.Collectively, these algorithms contribute to the sustainability and efficiency of cloud computing infrastructure. They provide a roadmap for researchers, practitioners, and decision-makers to make informed choices regarding virtual machine placement strategies. As cloud computing continues to evolve, further research into hybrid approaches and refinements will advance the field of energy-efficient VM placement, making cloud data centers more environmentally friendly and economically viable. In an era where the demand for computing resources is ever-increasing, energy-conscious solutions are essential to shape the future of cloud computing

REFERENCES

[1] Pavithra, B., & Ranjana, R. (2016). Energy-efficient resource provisioning with dynamic VM placement uses an energy-aware cloud load balancer. 2016 International Conference on Information Communication and Embedded Systems, ICICES 2016. https://doi.org/10.1109/ICICES.2016.7518919

[2] Zheng, X., & Cai, Y. (2015). Energy-Efficient Statistical Live Virtual Machine Placement for Big Data Information Systems in Cloud Computing Environments. 2015 IEEE International Conference on Smart City/SocialCom/SustainCom (SmartCity), 1053–1058. https://doi.org/10.1109/SmartCity.2015.208

[3] Barlaskar, E. (2014). Dynamic Consolidation of Virtual Machines with Multi-Agent System, 94(9), 30–38. Applications (0975 – 8887) Volume 55– No.13, October 2012.

[4] Susila, N. (2016). Energy Efficient Extended FCFS Load Balancing in Data Centers of Cloud, 11(1), 599–605.

[5] Xu, X., Cao, L., & Wang, X. (2016). Resource pre-allocation algorithms for low-energy task scheduling of cloud computing. Journal of Systems Engineering and Electronics, 27(2), 457–469. https://doi.org/10.1109/JSEE.2016.00047

[6] Seema Vahora, Ritesh Patel, Hetal Patel and Sandipkumar Patel"Efficient Virtual Machine Management Based on Dynamic Workload in Cloud Computing Environment"978-1-4799-8047-5/15/$31.00_c 2015 IEEE.

[7] Thorat, A., & Sonkar, P. S. (2016). Energy Aware Load Balancing Technique for Managing Network Workload in Cloud Computing, 4(3), 62–66.

[8] Gaggero, M., & Caviglione, L. (2016). Model Predictive Control for the Placement of Virtual Machines in Cloud Computing Applications, 1987–1992.

[9] Goyal, A., & Chahal, N. S. (2015). A Proposed Approach for Efficient Energy Utilization in Cloud Data Center, 115(11), 24–27.

[10] Bose, S., & Kumar, J. (2015). An Energy-Aware Cloud Load Balancing Technique using Dynamic Placement of Virtualized Resources, 2(7), 81–86.

[11] Chen, J. (2015). A Live Migration Algorithm for Virtual Machine in a Cloud Computing Environment. https://doi.org/10.1109/UIC-ATC-ScalCom-CBDCom-IoP.2015.239

[12] Bose, S., & Kumar, J. (2015). A Survey on Energy-Aware Load Balancing Techniques in Cloud Computing, 4(5), 397–402. https://doi.org/10.17148/IJARCCE.2015.4587

[13] Gandhi, N. K., & Kailu, P. (2015). Prediction Based Energy Efficient Virtual Machine Consolidation in Cloud Computing, 437–441. https://doi.org/10.1109/ICACCE.2015.148