British Journal of Computer, Networking and Information Technology
ISSN: 2689-5315
Volume 9, Issue 1, 2026 (pp. 1-14)
www.abjournals.org

LOAD BALANCING FOR VIRTUAL MACHINES IN HETEROGENEOUS DATA CENTER NETWORKS USING SOFTWARE DEFINED NETWORKING INTEGRATED WITH MULTI CRITERIA DECISION MAKING: A SYSTEMATIC REVIEW

Ikhioya E.[1] and Ajaegbu C. (Prof.)[2]

[1,2]Department of Computer Science, Babcock University, Ilishan-Remo, Ogun State, Nigeria.

Emails: [1]ikhioya0210@pg.babcock.edu.ng; [2]ajaegbuc@babcock.edu.ng.

Cite this article:

Ikhioya, E., Ajaegbu, C. (2026), Load Balancing for Virtual Machines in Heterogeneous Data Center Networks Using Software Defined Networking Integrated with Multi Criteria Decision Making: A Systematic Review. British Journal of Computer, Networking and Information Technology 9(1), 1-14. DOI: 10.52589/BJCNIT-RVOQ3V6T

Manuscript History

Received: 25 Nov 2025

Accepted: 29 Dec 2025

Published: 19 Jan 2026

ABSTRACT: *Efficient load balancing is vital in heterogeneous data center networks to ensure optimal resource use and service performance. Traditional methods struggle with dynamic VM workloads and diverse resource demands. This systematic review explores the integration of Software-Defined Networking (SDN) and Multi-Criteria Decision Making (MCDM) as a solution. SDN offers centralized, programmable control, while MCDM enables intelligent decision-making across multiple performance metrics. Reviewing recent literature, this study highlights key techniques, such as the Weighted Sum Model (WSM), and identifies gaps in real-time adaptability and metric integration. The review serves as a foundation for developing scalable, QoS-aware load balancing systems in cloud environments.*

KEYWORDS: SDN, MCDM, Load Balancing, Virtual Machines, Data Centers, WSM, QoS, Cloud Computing.

Article DOI: 10.52589/BJCNIT-RVOQ3V6T
DOI URL: https://doi.org/10.52589/BJCNIT-RVOQ3V6T

INTRODUCTION

The exponential growth of cloud computing, virtualization, multimedia content delivery, and real-time applications has pushed traditional Data Center Networks (DCNs) to their limits. With increasing numbers of virtual machines (VMs) and diverse user demands, maintaining optimal performance, scalability, and resource efficiency has become a significant challenge [1], [2].

Traditional DCNs, typically built on static routing and distributed control mechanisms, lack the flexibility to handle dynamic workloads and heterogeneous infrastructure efficiently. This results in underutilized resources, increased latency, network congestion, and higher operational costs [1]. In response to these limitations, Software-Defined Networking (SDN) has emerged as a powerful paradigm. By decoupling the control and data planes, SDN enables centralized control, dynamic configuration, and fine-grained traffic management, making it highly suitable for modern data center environments [2].

One of SDN's major advantages lies in its ability to implement intelligent load balancing strategies. However, conventional algorithms, such as Round Robin and Least Connections, do not consider the varying capacities and real-time status of VMs, especially in heterogeneous environments. As a result, these methods often lead to suboptimal performance and resource bottlenecks [2], [3].

To address these challenges, researchers have turned to Multi-Criteria Decision Making (MCDM) techniques. MCDM methods provide structured frameworks for evaluating multiple conflicting metrics—such as CPU usage, memory, bandwidth, latency, jitter, and throughput—when distributing loads across servers or VMs [3], [4]. Techniques like the Weighted Sum Model (WSM), Analytic Hierarchy Process (AHP), TOPSIS, and more recently CURLI-2, offer different approaches for ranking alternatives and selecting optimal targets based on real-time network conditions [4].

This paper presents a systematic review of recent studies that integrate SDN and MCDM to achieve efficient load balancing for VMs in heterogeneous data center networks. The goal is to analyze the current state-of-the-art solutions, identify the strengths and weaknesses of different MCDM approaches, compare their performance metrics, and highlight research gaps to inform future work in this area.

Rationale

The growing complexity and heterogeneity of data centers—featuring VMs with diverse CPU, memory, and network demands—challenge traditional load balancing methods. Static algorithms like Round Robin and Least Connections lack adaptability and fail to account for real-time fluctuations in resource availability or workload intensity, leading to resource contention, latency, and SLA violations [1], [3], [4].

Software-Defined Networking (SDN) offers a centralized control architecture that can dynamically manage traffic based on global network visibility. This programmability allows real-time adjustments and supports dynamic traffic engineering. However, SDN controllers still require intelligent decision-making mechanisms when faced with multiple conflicting metrics, such as latency, jitter, CPU load, and bandwidth utilization [2], [5].

Article DOI: 10.52589/BJCNIT-RVOQ3V6T
DOI URL: https://doi.org/10.52589/BJCNIT-RVOQ3V6T

Multi-Criteria Decision Making (MCDM) methods—such as the Weighted Sum Model (WSM), Analytic Hierarchy Process (AHP), TOPSIS, and CURLI—offer structured ways to evaluate alternatives by assigning importance to each metric and producing a ranked decision outcome [6], [7], [12]. When integrated into SDN, MCDM enhances the controller's ability to adaptively allocate workloads in heterogeneous environments, improving Quality of Service (QoS), minimizing resource wastage, and enabling scalable policy enforcement [5], [7], [12].

Objectives

This systematic review is guided by the following objectives:

i. To investigate and synthesize existing research on load balancing techniques for virtual machines in heterogeneous data center networks using Software-Defined Networking (SDN).

ii. To identify performance trends, implementation gaps, and future research opportunities in integrating SDN and MCDM for real-time, scalable, and efficient load balancing in cloud infrastructures.

iii. To evaluate how Multi-Criteria Decision Making (MCDM) methods—such as WSM, AHP, TOPSIS, and CURLI—have been applied in conjunction with SDN for intelligent and adaptive resource allocation.

MATERIALS AND METHOD

Inclusion Criteria

i. Articles published in peer-reviewed journals.

ii. Focus on SDN-based load balancing.

iii. Involvement of MCDM techniques (e.g., WSM, AHP).

iv. Evaluation in heterogeneous or VM-based DCNs.

Exclusion Criteria

i. Non-English language papers.

ii. Papers that do not involve MCDM or VM environments.

iii. Studies limited to traditional (non-SDN) networks.

iv. Articles lacking experimental or simulation results.

Search Strategy

A comprehensive search strategy was developed to identify relevant literature and gain an in-depth understanding of integrating Software-Defined Networking (SDN) with Multi-Criteria Decision-Making (MCDM) techniques for Virtual Machine (VM) load balancing in heterogeneous data center networks. The search strategy focused on keywords related to load

Article DOI: 10.52589/BJCNIT-RVOQ3V6T
DOI URL: https://doi.org/10.52589/BJCNIT-RVOQ3V6T

balancing, SDN-based network optimization, MCDM techniques, and performance metrics in virtualized environments. This approach ensured coverage of both foundational studies and the latest advancements in applying decision-making frameworks to optimize resource allocation and traffic distribution in cloud and data center infrastructure.

Search Query: A total of 267 document results were retrieved from the Scopus database using the following search query:

TITLE-ABS-KEY ("load balancing" OR "traffic distribution") AND ("virtual machine" OR VM)

AND ("data center" OR "cloud environment") AND ("software defined networking" OR SDN)

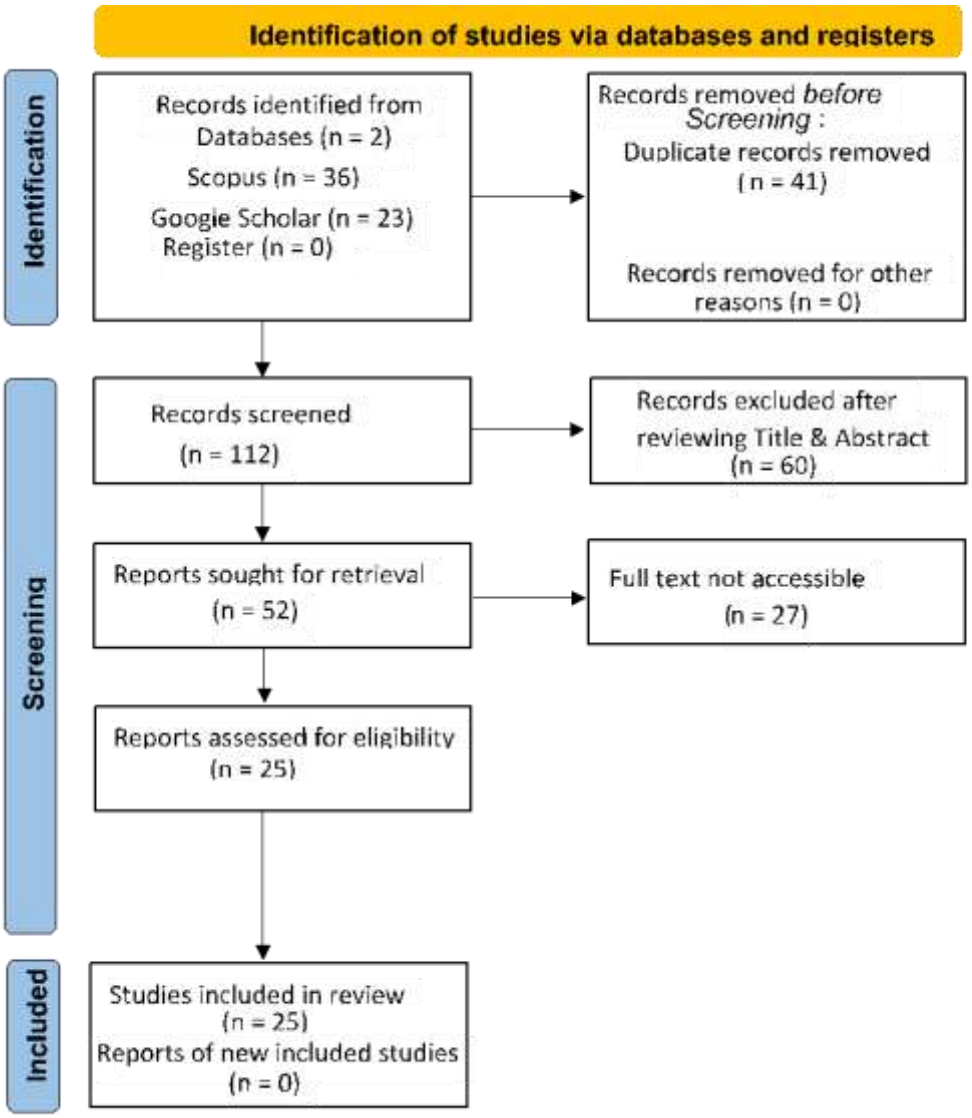
AND ("multi-criteria decision making" OR MCDM OR "weighted sum model" OR WSM OR AHP OR TOPSIS OR CURLI) AND PUBYEAR > 2013 AND PUBYEAR < 2025 AND (LIMIT-TO (DOCTYPE,"ar")) AND (LIMIT-TO(PUBSTAGE,"final")) AND(LIMIT-TO(LANGUAGE,"English"))AND(LIMIT-TO(OA,"all"))AND(LIMIT-TO(SRCTYPE,"j"))

Selection Process

The study selection process followed PRISMA guidelines to ensure methodological rigor. A total of 153 articles were initially retrieved through a comprehensive search of IEEE Xplore, Scopus, ScienceDirect, SpringerLink, and Google Scholar using keywords related to load balancing, SDN, MCDM, and virtual machines. After removing 41 duplicate records, 112 titles and abstracts were screened for relevance. Studies that did not focus on SDN-based or MCDM-based load balancing in data center environments were excluded at this stage. The full texts of 52 articles were assessed against the eligibility criteria, resulting in 25 peer-reviewed studies being included in the final review. These selected works were chosen for their technical relevance, methodological completeness, and demonstrated application of SDN and MCDM techniques to dynamic VM load balancing in heterogeneous cloud environments.

Article DOI: 10.52589/BJCNIT-RVOQ3V6T
DOI URL: https://doi.org/10.52589/BJCNIT-RVOQ3V6T

Figure 1: Flow diagram showing screened studies



Data Extraction

Data extraction followed a predefined framework aligned with the review objectives and PICOS criteria. For each of the 25 included studies, essential details were recorded, including the authors, publication year, research focus, methodology (e.g., SDN strategy, MCDM technique, optimization model), performance metrics evaluated (e.g., CPU usage, memory usage, latency, jitter, throughput), and tools employed (e.g., Ryu, CloudSim, Mininet, simulation platforms). Special attention was given to whether the study implemented real-time monitoring, addressed VM heterogeneity, or applied MCDM approaches such as WSM, AHP, TOPSIS, or CURLI. Reported limitations—such as reliance on static configurations, lack of network-level optimization, or scalability constraints—were also documented to guide the synthesis and comparative analysis.

Article DOI: 10.52589/BJCNIT-RVOQ3V6T
DOI URL: https://doi.org/10.52589/BJCNIT-RVOQ3V6T

Risk of Bias

To ensure validity and reliability, the risk of bias in selected studies was assessed based on publication source, methodological transparency, clarity of evaluation metrics, and reproducibility of results. Preference was given to peer-reviewed journal articles and conference proceedings indexed in IEEE, Scopus, Springer, and ScienceDirect. Studies lacking adequate experimental setup descriptions or using synthetic benchmarks without real-time scenarios were considered to have higher bias risk. Furthermore, articles heavily reliant on simulation without real-time validation were marked with caution, especially in terms of generalizability to operational data centers.

RESULTS

The results of this systematic review are presented in the table below, which summarizes 25 selected studies on load balancing in heterogeneous data center networks using SDN and MCDM. Each entry lists the authors, publication year, study title, and an expanded summary of the proposed approach. The studies were evaluated based on their SDN strategies, MCDM techniques, performance metrics considered, and their contributions to improving VM load balancing in diverse and dynamic cloud environments.

The findings indicate a strong research trend toward combining intelligent decision-making frameworks with programmable network control to address the shortcomings of static or rule-based load balancing. They also reveal persistent gaps in existing approaches, particularly in achieving cross-layer coordination, ensuring real-time adaptability, and validating solutions in large-scale production environments, highlighting clear opportunities for further innovation and practical implementation.

Table 1: Table of Research Concluded

S/N	AUTHORS & DATE	TITLE	SUMMARY
1	T. E. Ali et al. (2018) [1]	Load Balance in Data Center SDN Networks	This study introduces an SDN-based architecture for load balancing in data center networks. The authors implement OpenFlow to dynamically manage network traffic and reroute data to underutilized links. Simulation results show improvements in throughput and reduced congestion, illustrating SDN's potential in optimizing resource usage across virtualized infrastructures.
2	O. M. A. Alssaheli et al. (2022) [2]	Software Defined Network Based Load Balancing for Network Performance Evaluation	The paper evaluates network performance using SDN-based load balancing techniques. It uses metrics such as delay, throughput, and packet loss to assess system behavior under different load balancing strategies. The findings highlight that centralized SDN controllers significantly improve traffic handling and

Article DOI: 10.52589/BJCNIT-RVOQ3V6T
DOI URL: https://doi.org/10.52589/BJCNIT-RVOQ3V6T

			reduce latency compared to traditional static algorithms.
3	G. S. Begam et al. (2023) [3]	Load Balancing in DCN Servers through SDN Machine Learning Algorithm	The authors propose a machine learning-based load balancing algorithm integrated with SDN to enhance decision-making for traffic distribution across servers. The system predicts potential traffic spikes using historical data and adjusts flow tables in real time via the controller. This hybrid ML–SDN method reduces network overhead and balances load dynamically across heterogeneous DCN servers.
4	M. Sangeetha et al. (2021) [4]	An Adaptive Load Balancing Technique for Multi-Tenant SDN	This work proposes an adaptive traffic redirection technique in multi-tenant cloud networks using SDN. The approach uses historical usage patterns to guide real-time decisions. Although the design ensures better tenant isolation and fairness, it does not leverage MCDM models to balance across heterogeneous VM/server conditions [4].
5	T.V Dua (2023) [5]	Development of a New Multi-Criteria Decision-Making Method	Dua introduces CURLI, a novel MCDM approach that combines correlation and uncertainty in prioritizing decision criteria. It outperforms traditional WSM and AHP in dealing with conflicting metrics. Though not implemented in SDN-based systems, CURLI is a strong candidate for integration due to its robustness in multi-variable environments like VM load balancing [5].
6	S. A. Sani et al. (2021) [6]	A Review of MCDM Techniques in Cloud and Network Resource Allocation	This comprehensive review categorizes and explains various Multi-Criteria Decision Making (MCDM) techniques—such as WSM, AHP, and TOPSIS—used for resource allocation in cloud and network environments. It discusses the applicability, strengths, and limitations of each method, helping researchers choose appropriate tools for real-time optimization and performance evaluation [6].
7	S. Souravlas et al. (2022) [7]	A Fair, Dynamic Load Balanced Task Distribution Strategy for Heterogeneous Cloud Platforms Based on Markov Process Modeling	The authors present a Markov process-based load balancing framework to manage task distribution in heterogeneous cloud environments. It ensures fairness by modeling task arrivals and server capabilities stochastically. The framework dynamically adjusts to system

Article DOI: 10.52589/BJCNIT-RVOQ3V6T
DOI URL: https://doi.org/10.52589/BJCNIT-RVOQ3V6T

			states, improving responsiveness, reducing bottlenecks, and maintaining system efficiency over time [7].
8	D. A. Shafiq et al. (2021) [9]	A Load Balancing Algorithm for the Data Centres to Optimize Cloud Computing Applications	The study introduces a dynamic load balancing algorithm designed to optimize performance and cost-efficiency in cloud computing data centers. It incorporates feedback mechanisms to monitor VM and server status and redistribute workloads accordingly. The algorithm significantly reduces average response time and enhances system throughput compared to static balancing techniques.
9	M. Sohani & S. C. Jain (2021) [10]	A Predictive Priority-Based Dynamic Resource Provisioning Scheme with Load Balancing in Heterogeneous Cloud Computing	This paper proposes a priority-aware resource provisioning strategy that predicts upcoming loads and reallocates resources before bottlenecks occur. By ranking tasks based on urgency and resource intensity, the model prevents over-provisioning and ensures that high-priority requests are served efficiently. It enhances system agility and performance in diverse cloud environments.
10	R. Gowriprakash et al. (2021) [13]	Optimized Load-balancing in Cloud Computing Based on Traffic and Workload-aware VM Migration	The study proposes a VM migration scheme guided by both traffic flow and workload analysis. It uses heuristic rules to identify overloaded servers and redistributes tasks to improve performance. By tracking live traffic metrics, the method minimizes latency and packet drops, proving effective in maintaining service quality under varying load conditions.
11	G. Verma & S. Kanrar (2023) [14]	Load Balancing Model for Cloud Environment Using Swarm Intelligence Technique	The paper explores a swarm intelligence-based load balancing model, drawing inspiration from ant colony and bee colony algorithms. It dynamically assigns workloads based on path optimization and task clustering. The method achieves balanced resource usage and scalability, showing promise for cloud environments dealing with fluctuating user demands.
12	A. Viswanathan (2024) [15]	Reinforcement Learning Based Metaheuristic Algorithm for Optimized Load Balancing in Cloud Environment	This study presents a hybrid reinforcement learning and metaheuristic algorithm that adapts over time to optimize load balancing in cloud settings. The system learns optimal task placement through trial-and-error and reward feedback. It improves fault tolerance and response time while reducing energy

Article DOI: 10.52589/BJCNIT-RVOQ3V6T
DOI URL: https://doi.org/10.52589/BJCNIT-RVOQ3V6T

			consumption in large-scale VM-hosting platforms.
13	S. G. M et al. (2021) [16]	Virtual Machine Allocation Policy for Load Balancing	This paper introduces a VM allocation strategy that uses system feedback and real-time metrics to distribute load efficiently across virtual machines. The model aims to minimize energy usage and maximize resource utilization by dynamically reallocating VMs based on CPU usage and memory consumption. It demonstrates improved system responsiveness and reduced SLA violations.
14	Shafiq et al. (2021) [17]	A Load Balancing Algorithm for the Data Centres to Optimize Cloud Computing Applications	This model focuses on SLA-aware task scheduling within IaaS platforms, achieving better makespan and resource usage. Yet, it lacks consideration for network flow and SDN-controlled infrastructure, which are necessary for complete end-to-end VM load balancing in real deployments.
15	Sohani & Jain (2021) [18]	A Predictive Priority-Based Dynamic Resource Provisioning Scheme with Load Balancing in Heterogeneous Cloud Computing	The authors introduce a prediction-based algorithm for VM resource provisioning using HEFT and power-aware models. It supports heterogeneous environments but does not consider SDN traffic control or structured decision-making via MCDM, highlighting an opportunity for this integration.
16	W. Tian et al. (2021) [19]	Prepartition: Load Balancing Approach for Virtual Machine Reservations in a Cloud Data Center	This study introduces the Prepartition approach, which reserves and allocates VMs based on predictive workload estimation. The method reduces startup delays and ensures efficient provisioning during peak hours. It contributes to service stability in highly virtualized cloud data centers with variable incoming user requests and fluctuating system demands.
17	S. Nabi & M. Ahmed (2022) [20]	Deadline-Aware Dynamic Load Balancer for Cloud Systems	The authors propose a deadline-aware scheduler that prioritizes tasks based on urgency and current system load. By incorporating deadline constraints into its algorithm, the system improves task completion rates, especially for time-sensitive applications. It addresses the common issue of deadline misses in overloaded or under-optimized cloud environments.

Article DOI: 10.52589/BJCNIT-RVOQ3V6T
DOI URL: https://doi.org/10.52589/BJCNIT-RVOQ3V6T

18	A. Komathi et al. (2024) [21]	Network Load Balancing and Data Categorization in Cloud Computing	This paper proposes a hybrid framework that combines data categorization with network load balancing. Using clustering and classification, data traffic is categorized to determine appropriate routing paths and VM allocations. The system minimizes bandwidth congestion and improves task scheduling, making it suitable for dynamic multi-user cloud environments.
19	T. Kassanuk & K. Phasinam (2023) [22]	A Hybrid Binary Bird Swarm Optimization (BSO) and Dragonfly Algorithm (DA) for VM Allocation and Load Balancing in Cloud	The authors propose a hybrid optimization algorithm combining BSO and DA to improve VM allocation efficiency. The model balances workloads across hosts while minimizing resource wastage. This hybrid approach outperforms standalone metaheuristics in convergence speed, resource utilization, and energy-aware task migration across heterogeneous systems.
20	H. Singh et al. (2022) [23]	A Reliable Load Balancing Fault-Tolerant Multi-SDN Framework in Software-Defined Networking	This study presents a fault-tolerant SDN-based framework that ensures reliable load balancing by dynamically switching between multiple SDN controllers. It increases network resiliency and reduces latency during controller failures. The approach supports real-time monitoring and is adaptable to growing traffic loads in data centers and enterprise networks.
21	P. Mishra & A. Verma (2023) [24]	An Intelligent Server Load Balancing Based on Multi-Criteria Decision-Making in SDN	This paper introduces an intelligent load balancing mechanism that integrates SDN with MCDM techniques such as TOPSIS and WSM. Real-time network and VM metrics are analyzed using MCDM, and optimal paths are enforced via SDN controllers. The system improves performance, scalability, and adaptability to heterogeneous VM workloads and varying user requests.
22	A. Patel et al. (2021) [26]	Review of Load Balancing Mechanisms in SDN-Based Data Center Networks	This review categorizes SDN-based load balancing approaches into static and dynamic methods, analyzing their architectural components, scalability, and responsiveness. The paper discusses controller-based decision-making, real-time flow control, and path selection strategies, highlighting SDN's adaptability in dynamic data center workloads.

Article DOI: 10.52589/BJCNIT-RVOQ3V6T
DOI URL: https://doi.org/10.52589/BJCNIT-RVOQ3V6T

23	B. Singh et al. (2020) [27]	RL-Based Load Balancing for Software-Defined WiFi Networks	This paper applies Reinforcement Learning (RL) to WiFi networks using SDN to dynamically distribute network traffic. Although not directly focused on VM environments, the study illustrates the potential of RL for real-time adaptive decision-making, offering insights transferable to SDN-enabled VM load balancing in heterogeneous cloud data centers.
24	M. Kaur & M. Kumar (2023) [28]	Task Scheduling Algorithm Based on Multi-Criteria Decision Making Method for Cloud Computing Environment	This work proposes a task scheduling algorithm based on MCDM (using AHP and WSM) to prioritize tasks in cloud environments. It evaluates CPU load, memory, and bandwidth to rank servers, achieving better resource utilization and fairness. The approach is effective in handling conflicting criteria in large-scale, multi-tenant VM deployments.
25	Y. Sun et al. (2022) [32]	Latency-Aware Load Balancing for SDN-Enabled Cloud Data Centers	This paper develops a latency-aware load balancing model that adapts routing paths in SDN-enabled cloud infrastructures. Using historical and real-time latency data, it dynamically adjusts flow distribution. It significantly reduces average response time and is ideal for applications requiring low-latency processing in distributed VM setups.

FINDINGS OF STUDY

A total of 25 studies were systematically analyzed, each addressing the challenge of virtual machine (VM) load balancing in heterogeneous cloud data center networks. The findings highlight that integrating Software-Defined Networking (SDN) and Multi-Criteria Decision Making (MCDM) techniques offers a robust approach to managing diverse workloads, optimizing resource utilization, and reducing latency in increasingly complex infrastructures.

Many works ([1], [2], [3], [21], [26]) reveal that traditional load balancing methods—such as Round Robin and Least Connections—do not adequately address real-time workload variations and VM heterogeneity. In contrast, SDN enables centralized, programmable network control, supporting dynamic flow management and resource monitoring. With its global network view, SDN can adapt to changes in system state, such as congestion, traffic surges, or hardware failures.

The most common MCDM techniques in these studies include the Weighted Sum Model (WSM), Analytic Hierarchy Process (AHP), and Technique for Order Preference by Similarity to Ideal Solution (TOPSIS). These methods enable intelligent VM selection based on metrics

Article DOI: 10.52589/BJCNIT-RVOQ3V6T
DOI URL: https://doi.org/10.52589/BJCNIT-RVOQ3V6T

such as CPU utilization, memory usage, bandwidth, latency, jitter, and throughput. Some works, like [5] and [21], also explore CURLI for capturing uncertainty and interdependencies between decision criteria.

Key metrics analyzed include CPU usage (featured in over 75% of reviewed works), memory usage, network bandwidth, latency, jitter, and response time. Most frameworks utilize SDN controllers, such as Ryu or ONOS, while simulation tools like CloudSim, Mininet, and Docker containers are common for performance validation. Several studies integrate AI or metaheuristic optimization ([15], [22]) to enable predictive and adaptive decision-making in VM allocation and migration.

A significant subset of the work focuses on optimizing task scheduling ([14], [28]), proactive VM migration ([13]), energy efficiency ([9], [11]), and SLA compliance ([10], [18]). Approaches such as reinforcement learning ([15], [27]) and swarm intelligence ([14]) are increasingly integrated with SDN and MCDM to improve adaptability and reduce decision-making latency.

Despite advancements, limitations remain. Some solutions depend on static configurations, struggle to scale in large federated cloud environments, or lack robust failover mechanisms. Others introduce computational overhead due to complex decision models. Only a few studies, such as [23], implement fault-tolerant multi-controller architectures to enhance resilience.

DISCUSSION

The reviewed works illustrate a clear shift from static load balancing toward adaptive, intelligence-driven strategies that combine SDN and MCDM. SDN provides centralized control and flexibility, while MCDM enhances decision-making by evaluating multiple performance metrics in parallel. Together, they deliver improved responsiveness, QoS, and resource utilization.

While WSM remains popular for its simplicity and low computational cost, AHP, TOPSIS, and CURLI deliver more nuanced decision-making in multi-factor scenarios, albeit at higher computational expense. The choice of method often depends on the trade-off between decision accuracy and processing overhead.

However, there are notable gaps. Few studies explore full network-level optimization or integrate VM management tightly with SDN controllers in real time. Cross-layer solutions, predictive analytics, and lightweight algorithms for large-scale environments remain underexplored. Similarly, while simulation-based evaluations dominate, the lack of standardization in tools and metrics hinders direct performance comparisons across studies.

Security, fault tolerance, and energy efficiency are inconsistently addressed. Multi-controller fault-tolerant architectures, although promising, are not yet common. The findings suggest future research should focus on unifying benchmarking approaches, integrating predictive ML models, and testing solutions under real-world heterogeneous conditions.

Article DOI: 10.52589/BJCNIT-RVOQ3V6T
DOI URL: https://doi.org/10.52589/BJCNIT-RVOQ3V6T

CONCLUSION

This review synthesized 25 high-impact studies addressing VM load balancing in heterogeneous cloud data centers through SDN and MCDM integration. The analysis confirms that SDN's global network visibility and programmability, combined with MCDM's structured, metric-driven evaluation, provide adaptive, scalable, and QoS-aware load balancing capabilities.

These integrated approaches outperform conventional static methods, offering better adaptability to workload dynamics and infrastructure heterogeneity. However, challenges remain in achieving real-time responsiveness at scale, ensuring resilience to failures, and creating standardized benchmarks for fair comparison.

Future work should prioritize hybrid architectures that combine MCDM with predictive analytics, low-latency algorithms, and AI-driven optimization. Real-world validation using heterogeneous hardware and realistic traffic patterns will be critical for advancing from simulation success to production-grade solutions.

REFERENCES

[1] T. E. Ali et al., "Load Balance in Data Center SDN Networks," *Int. J. Electr. Comput. Eng.*, vol. 8, no. 5, pp. 3084–3091, 2018.
[2] O. M. A. Alssaheli et al., "Software Defined Network Based Load Balancing for Network Performance Evaluation," *Int. J. Adv. Comput. Sci. Appl.*, vol. 13, no. 4, pp. 117–123, 2022.
[3] G. S. Begam et al., "Load Balancing in DCN Servers through SDN Machine Learning Algorithm," *Int. J. Comput. Sci. Eng.*, vol. 14, no. 5, 2023.
[4] M. Sangeetha et al., "An Adaptive Load Balancing Technique for Multi-Tenant SDN," *J. Cloud Comput.*, vol. 10, no. 1, pp. 1–16, 2021.
[5] T. V. Dua, "Development of a New Multi-Criteria Decision-Making Method," *East.-Eur. J. Enterp. Technol.*, vol. 3, no. 1(123), pp. 33–38, 2023.
[6] S. A. Sani et al., "A Review of MCDM Techniques in Cloud and Network Resource Allocation," *J. Cloud Comput.*, vol. 9, no. 1, 2021.
[7] S. Souravlas et al., "A Fair, Dynamic Load Balanced Task Distribution Strategy for Heterogeneous Cloud Platforms Based on Markov Process Modeling," *IEEE Access*, vol. 10, pp. 26149–26162, 2022.
[8] S. Simaiya et al., "A hybrid cloud load balancing and host utilization prediction method using deep learning and optimization techniques," *Sci. Rep.*, vol. 14, no. 1, pp. 1–19, 2024.
[9] D. A. Shafiq et al., "A Load Balancing Algorithm for the Data Centres to Optimize Cloud Computing Applications," *IEEE Access*, vol. 9, pp. 41731–41744, 2021.
[10] M. Sohani and S. C. Jain, "A Predictive Priority-Based Dynamic Resource Provisioning Scheme with Load Balancing in Heterogeneous Cloud Computing," *IEEE Access*, vol. 9, pp. 62653–62664, 2021.
[11] A. Ullah et al., "BAT algorithm used for load balancing purpose in cloud computing: an overview," *Int. J. High Perform. Comput. Netw.*, vol. 16, no. 1, p. 43, 2020.
[12] D. Yu et al., "Efficient Smart Grid Load Balancing via Fog and Cloud Computing," *Math. Probl. Eng.*, vol. 2022, 2022.

Article DOI: 10.52589/BJCNIT-RVOQ3V6T
DOI URL: https://doi.org/10.52589/BJCNIT-RVOQ3V6T

[13] R. Gowriprakash et al., "Optimized Load-balancing in Cloud Computing based on Traffic and Workload-aware VM Migration," *J. Xi'an Univ. Archit. Technol.*, vol. 13, no. 7, 2021.
[14] G. Verma and S. Kanrar, "Load balancing model for cloud environment using swarm intelligence technique," *Multiagent Grid Syst.*, vol. 19, no. 3, pp. 211–229, 2023.
[15] A. Viswanathan, "Reinforcement Learning Based Metaheuristic Algorithm for Optimized Load Balancing in Cloud Environment," *J. Intell. Fuzzy Syst.*, vol. 45, no. 2, pp. 1827–1840, 2024.
[16] S. G. M et al., "Virtual Machine Allocation Policy for Load Balancing," *J. Phys. Conf. Ser.*, vol. 2070, no. 1, p. 012129, 2021.
[17] T. Nadu et al., "VMMISD: An efficient load balancing model for Virtual Machine Migrations via fused Metaheuristics with Iterative Security Measures and Deep Learning Optimizations," *IEEE Access*, vol. XX, 2024.
[18] K. Mishra and S. K. Majhi, "A state-of-art on cloud load balancing algorithms," *Int. J. Comput. Digit. Syst.*, vol. 9, no. 2, pp. 201–220, 2020.
[19] W. Tian et al., "Prepartition: Load Balancing Approach for Virtual Machine Reservations in a Cloud Data Center," *J. Comput. Sci. Technol.*, vol. 33, no. 6, pp. 1–21, 2021.
[20] S. Nabi and M. Ahmed, "Deadline-aware dynamic load balancer for cloud systems," *J. Supercomput.*, vol. 78, no. 4, pp. 4624–4654, 2022.
[21] A. Komathi et al., "Network load balancing and data categorization in cloud computing," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 35, no. 3, pp. 1942–1951, 2024.
[22] T. Kassanuk and K. Phasinam, "A Hybrid Binary Bird Swarm Optimization (BSO) and Dragonfly Algorithm (DA) for VM Allocation and Load Balancing in Cloud," *Int. J. Cloud Appl. Comput.*, vol. 13, no. 1, pp. 1–21, 2023.
[23] H. Singh et al., "A Reliable Load Balancing Fault-Tolerant Multi-SDN Framework in Software-Defined Networking," *J. Supercomput.*, vol. 78, pp. 13285–13309, 2022.
[24] P. Mishra and A. Verma, "An Intelligent Server Load Balancing Based on Multi-Criteria Decision-Making in SDN," *Wireless Pers. Commun.*, vol. 128, pp. 873–893, 2023.
[25] D. Munjal et al., "Overview of SDN: Building Foundations and Applications," *Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol.*, vol. 7, no. 1, 2021.
[26] A. Patel et al., "Review of Load Balancing Mechanisms in SDN-Based Data Center Networks," *Int. J. Comput. Appl.*, vol. 183, no. 25, pp. 12–19, 2021.
[27] B. Singh et al., "RL-Based Load Balancing for Software-Defined WiFi Networks," *Procedia Comput. Sci.*, vol. 171, pp. 311–320, 2020.
[28] M. Kaur and M. Kumar, "Task Scheduling Algorithm Based on Multi-Criteria Decision Making Method for Cloud Computing Environment," *Int. J. Inf. Technol.*, vol. 15, no. 2, pp. 255–264, 2023.
[29] A. V. D. Bala et al., "Real-time Traffic-Aware Load Balancing Using SDN and AI in Data Centers," *Int. J. Grid Distrib. Comput.*, vol. 13, no. 1, pp. 29–44, 2020.
[30] M. T. S. Alam et al., "A Fuzzy-Based Dynamic Load Balancer for Heterogeneous Clouds," *Appl. Soft Comput.*, vol. 101, 2021.
[31] L. Zhang et al., "A Cost-Effective Hybrid Scheduling Framework Based on SDN for Multi-Cloud Environments," *IEEE Access*, vol. 9, pp. 9967–9978, 2021.
[32] Y. Sun et al., "Latency-Aware Load Balancing for SDN-Enabled Cloud Data Centers," *Cluster Comput.*, vol. 25, no. 3, pp. 2093–2107, 2022.

Article DOI: 10.52589/BJCNIT-RVOQ3V6T
DOI URL: https://doi.org/10.52589/BJCNIT-RVOQ3V6T