Nanotechnology Perceptions
ISSN 1660-6795
www.nano-ntp.com

Optimized Load Balancing Techniques in Cloud Computing Environment: A Systematic Literature Review and Future Trends

Sonal Shah[1], Dr. Papri Das[2], Dr. Akhilesh Latoria[3], Dhaval J. Thaker[4]

[1]*Research scholar, School of Information Technology, AURO University, Surat, Gujarat, India*
[2]*Associate Professor, School of Information Technology, AURO University, Surat, Gujarat, India*
[3]*Associate Professor, School of Engineering and Sciences, GD Goenka University, Sohna Gurugram, Haryana, India*
[4]*Assistant Professor, School of Information Technology, AURO University, Surat, Gujarat, India*

In today's rapidly evolving computing landscape, cloud computing has emerged as a pivotal paradigm that offers scalable and flexible resource provisioning to meet the demands of diverse applications. A critical aspect in this environment is load balancing, a technique that optimally distributes workloads across available resources to ensure efficient resource utilization, enhanced system performance, and seamless user experiences. This systematic literature review (SLR) delves into the realm of optimized load balancing techniques within cloud computing environments, encompassing various approaches, optimization methods, challenges, and future trends. The SLR employs a structured methodology to curate and analyze a comprehensive collection of 50 research articles. Through this examination, the evolution of load balancing techniques is traced from traditional approaches to the more dynamic and adaptive strategies that characterize contemporary cloud infrastructures. By categorizing these techniques into centralized, decentralized, and hybrid methodologies, the review illuminates the distinctive features and limitations of each approach. Optimization methods play a pivotal role in refining load balancing mechanisms. These methods encompass a spectrum of strategies, including meta-heuristics, machine learning, and swarm intelligence, all aimed at achieving optimal allocation of resources and minimizing response times. The SLR underscores the significance of optimization in addressing challenges posed by heterogeneity, scalability, and real-time adaptability. This review also identifies imminent challenges faced by load balancing in cloud computing environments. These challenges encompass the intricate interplay between resource allocation, fluctuating workloads, and dynamic system conditions. Additionally, the review offers insights into future trends that include harnessing the power of artificial intelligence and machine learning techniques, exploring fog and edge computing for load distribution, and exploring hybrid load balancing strategies to further enhance system efficiency. The findings of this systematic literature review contribute to a holistic understanding of optimized load balancing in cloud computing environments. By encapsulating the evolution of techniques, the role of optimization methods, and the challenges faced, the review provides a comprehensive foundation for researchers and practitioners alike. Moreover, the review sets the stage for future investigations by highlighting emerging trends that hold the potential to shape the future of load balancing in cloud computing. In the subsequent sections, we delve deeper into the current state of research, outlining the proposed approach in alignment with the research objectives. Subsequently, we expound upon the experimental outcomes attained through the proposed framework, encapsulating both its methodology and results.

Keywords: Dynamic load balancing, reactive fault tolerance, cloud, systematic literature review, resource allocation, task scheduling, virtual machine, workload management, optimization.

1. Introduction

Load balancing is a critical aspect of cloud computing, aimed at distributing computational tasks evenly across available resources to enhance system performance, resource utilization, and user satisfaction. This systematic literature review aims to explore the existing literature on optimized load balancing techniques in the cloud computing environment. The review analyses various approaches, methodologies, and trends while identifying challenges and suggesting potential future directions.

The contemporary landscape of information technology has been revolutionized by the advent of cloud computing, a transformative paradigm that underpins a multitude of applications across diverse sectors, including business operations and scientific research [14, 46]. This monumental advancement has ushered in unparalleled levels of scalability, agility, and cost-effectiveness, facilitated by the provision of on-demand access to computing resources [44, 38]. However, this rapid evolution has not been without its challenges, and within the realm of cloud computing, efficient resource allocation and management through load balancing have emerged as paramount concerns [29, 26].

Load balancing, a fundamental technique deeply entrenched in the domain of cloud computing, serves as the linchpin for optimizing resource utilization and ensuring resilient system performance [16, 15]. At its core, load balancing entails the equitable distribution of workloads across available resources to mitigate issues of resource underutilization or overutilization, both of which can culminate in performance bottlenecks and exacerbated response times [4, 24]. The cardinal aim of this technique is to harmonize the distribution of computational tasks, thereby bolstering system efficiency and responsiveness.

However, the pursuit of optimal load distribution is fraught with multifaceted challenges, primarily stemming from the dynamic and unpredictable nature of cloud environments [8, 21]. These environments are characterized by the ebb and flow of workloads, resource heterogeneity, and real-time variations in demand [9, 3]. As such, achieving an optimal equilibrium in the allocation of computational tasks becomes a complex endeavor, necessitating the synthesis of innovative strategies to tackle these challenges.

The systematic literature review (SLR) embarked upon in this research seeks to provide a comprehensive exploration of the landscape of optimized load balancing techniques within cloud computing environments. Through this SLR, we aim to unravel the evolutionary trajectory of load balancing methodologies, scrutinize the pivotal role of optimization techniques in enhancing load balancing efficacy, highlight the intricacies of challenges faced, and project emerging trends that will shape the trajectory of this field [11, 25].

Amidst the dynamic backdrop of the ever-evolving cloud computing domain, our research endeavors to harmonize insights gleaned from a rich tapestry of scholarly works [47, 8]. To accomplish this, a meticulous selection of 50 research articles has been meticulously analyzed, thereby furnishing a holistic view of the multifaceted landscape of optimized load balancing within cloud environments.

The multifaceted problem at hand encapsulates the nuances of resource allocation, task distribution, and performance optimization, rendering it a challenge that necessitates nuanced approaches [19, 11]. In light of the expanding cloud ecosystem, the exigency of identifying strategies that efficaciously manage the intricate interplay of workloads across distributed resources is paramount [26, 29]. Consequently, our research delves into the realm of optimized load balancing techniques, with the overarching goal of illuminating the methodologies that have been proffered, the optimization paradigms they harness, the intricacies of the challenges they grapple with, and the trajectories they chart for future progress.

Furthermore, our study forges an innovative path by systematically analyzing the feature extraction processes that underpin the examined literature [15, 8]. Feature extraction, a pivotal step in this research, unveils the defining characteristics of load balancing techniques and their performance attributes. By assimilating these features, we aspire to furnish a comprehensive understanding of the merits and limitations of disparate load balancing approaches.

Consequently, this research study unfolds as a conduit for identifying emerging trends and paradigms that hold the potential to sculpt the trajectory of optimized load balancing in the realm of cloud computing [14, 26]. Insights gleaned from this study transcend mere understanding, propelling us toward a vantage point from which we can embark on more informed and innovative trajectories, thus enhancing the efficiency and efficacy of resource allocation within the dynamic fabric of cloud environments.

In the subsequent sections, the intricate tapestry of background information will be woven, effectively contextualizing the research, elucidating the precise nature of the problem, outlining the study's overarching purpose, shedding light on the feature extraction process, and finally presenting the cardinal contributions that this research endeavor bequeaths to the domain of optimized load balancing in cloud computing environments.

2. Systematic literature review

The review involved analyzing 50 research papers that span across different journals and conference proceedings. These papers were selected based on their relevance to the topic of optimized load balancing techniques in cloud computing environments. The papers were categorized based on the load balancing approaches, techniques, algorithms, and their applications.

In the pursuit of enhancing load balancing in cloud computing environments, researchers have developed a multitude of techniques, each designed to address specific challenges posed by the dynamic and heterogeneous nature of such environments. In this section, we delve into a comprehensive examination of existing techniques related to the recommended study, outlining their benefits, drawbacks, and the distinct voids they leave unfilled. Through this scrutiny, we underscore the necessity of proposing a novel system that can surmount the limitations of current approaches and set new benchmarks for load balancing efficacy.

2.1 Centralized load balancing techniques

Centralized load balancing techniques leverage a central node to manage task allocation and resource distribution [2, 3]. These techniques offer a simplified approach to load management, allowing for a unified control mechanism. However, they often encounter scalability challenges as the system expands, leading to potential bottlenecks and single points of failure [1, 5]. Moreover, they might not adequately cater to the real-time variations in workloads characteristic of cloud environments [8]. As such, while centralized techniques provide a foundational understanding of load balancing, they fall short in accommodating the complexities inherent in modern cloud ecosystems.

2.2 Decentralized load balancing techniques

Decentralized techniques distribute load balancing decisions across various nodes, enhancing scalability and fault tolerance [4, 19]. These techniques exploit the autonomy of distributed nodes, enabling responsive load management in dynamic environments. However, they can suffer from synchronization and coordination issues, potentially leading to load imbalances [8, 3]. Additionally, these techniques might not achieve the same level of system-wide optimization as centralized approaches.

2.3 Hybrid load balancing techniques

Recognizing the strengths of both centralized and decentralized methods, hybrid techniques amalgamate their benefits [4, 3]. By striking a balance between centralized control and distributed autonomy, these techniques attempt to mitigate the shortcomings of individual approaches. While offering enhanced scalability and responsiveness, hybrid methods can also introduce increased complexity in terms of implementation and management [8, 29]. Additionally, optimizing the interplay between central and distributed nodes remains a challenge.

2.4 Meta-heuristic optimization techniques

Meta-heuristic optimization methods, including particle swarm optimization (PSO), genetic algorithms, and ant colony optimization, infuse load balancing strategies with adaptive and learning capabilities [19, 20]. These techniques dynamically adapt to changing workloads, allowing for effective resource allocation in real-time. However, they might require substantial computational overhead and parameter tuning, rendering them less suitable for environments with stringent resource constraints [9, 18].

2.5 Machine learning-based load balancing techniques

Leveraging the power of machine learning, these techniques employ algorithms to learn from historical workload patterns and optimize resource allocation [26, 15]. Such approaches hold the promise of self-adaptation and improved prediction accuracy. Nonetheless, they necessitate extensive training data and might not be suitable for highly dynamic environments with rapidly changing workloads [8, 26].

2.6 Identified gaps and novelty of the proposed system

Despite the notable progress made in load balancing techniques, several gaps persist in the current research landscape. The majority of existing approaches focus on optimizing load distribution within cloud environments, but they often neglect the nuances introduced by emerging paradigms such as fog and edge computing [15, 11]. Moreover, while many techniques target resource allocation and response time optimization, fewer emphasize energy efficiency and environmental sustainability [29, 19]. Furthermore, scalability remains a recurrent challenge across various techniques, and few address the complex trade-offs between scalability and fault tolerance [8, 15]. Here's a table summarizing the relevant load balancing techniques discussed in the previous section, along with their benefits, drawbacks, and identified gaps:

Table 1 Summarizing of the relevant load balancing techniques

Technique Category	Techniques	Benefits	Drawbacks	Identified Gaps
Centralized Load Balancing	Milani & Navimipour, 2016 [1]	Simplified control	Scalability issues	Lack of real-time adaptability
	Ghomi et al., 2017 [3]	Unified management	Potential single point of failure	
Decentralized Load Balancing	Kumar & Kumar, 2019 [4]	Scalability	Synchronization issues	Lack of system-wide optimization
	Alghamdi, 2022 [19]	Fault tolerance	Load imbalances	
Hybrid Load Balancing	Thakur & Goraya, 2017 [8]	Scalability & responsiveness	Implementation complexity	Optimizing central-distributed interplay
	Panwar et al., 2022 [29]	Balance between control & autonomy		
Meta-Heuristic Optimization	Devaraj et al., 2020 [9]	Real-time adaptation	Computational overhead	Resource constraints
	Shahid et al., 2020 [18]	Adaptive optimization	Parameter tuning	
Machine Learning-based	Gures et al., 2022 [26]	Self-adaptation & prediction accuracy	Extensive training data	Rapidly changing workloads
	Kaur & Aron, 2021 [15]	Learning from historical patterns	Limited suitability for dynamic environments	

This table provides a summarized overview of the discussed techniques, their benefits, drawbacks, and identified gaps. It serves as a quick reference for understanding the landscape of load balancing techniques in cloud computing environments.

The proposed system, with its focus on addressing these gaps, emerges as a novel contribution to the field of optimized load balancing in cloud computing environments. By integrating emerging paradigms, such as fog and edge computing, and prioritizing energy-efficient load distribution, our system aims to establish a holistic approach that bridges existing gaps and paves the way for more comprehensive load balancing solutions.

In the forthcoming sections, we delve into the architecture and design of our proposed system, elucidating the methodologies that underpin its development and highlighting its unique contributions to overcoming the limitations of current load balancing techniques.

3. Needs for systematic literature review

3.1 Methodology: systematic literature review

This section presents the methodological framework employed for conducting the systematic literature review (SLR), encompassing the review protocol and the meticulously devised search strategy. The systematic review methodology facilitated the comprehensive exploration of optimized load balancing techniques in cloud computing environments, ensuring a thorough and unbiased synthesis of existing knowledge.

3.2 Review protocol and search strategy

The review protocol was meticulously designed to encompass a wide spectrum of relevant literature. The investigation focused on dynamic load balancing, reactive fault tolerance, and related optimization techniques in cloud computing environments. The research was conducted using databases renowned for scholarly literature, including Google Scholar, IEEE Xplore, Scopus, and PubMed.

3.3 Keywords and search terms

The search was orchestrated using a spectrum of pertinent keywords to ensure a comprehensive exploration. The identified keywords and search terms included:

Load balancing techniques in cloud computing environment.

An efficient load balancing approach in cloud computing environment.

Load balancing in cloud environment.

Optimized Approach for Load Balancing in Cloud.

Analysis of Load Balancing Used in Cloud Computing Environment.

Cloud computing based on load balancing techniques.

Article Identification and Screening:

The search yielded a substantial collection of articles that matched the designated criteria. The following number of text articles were retrieved from the specified databases within the given timeframe:

Google Scholar: 1500 articles (2018-2023)

IEEE Xplore: 250 articles (using search term 5)

Scopus: 800 articles (using search terms 1-6)

PubMed: 65 articles (using search terms 1-6)

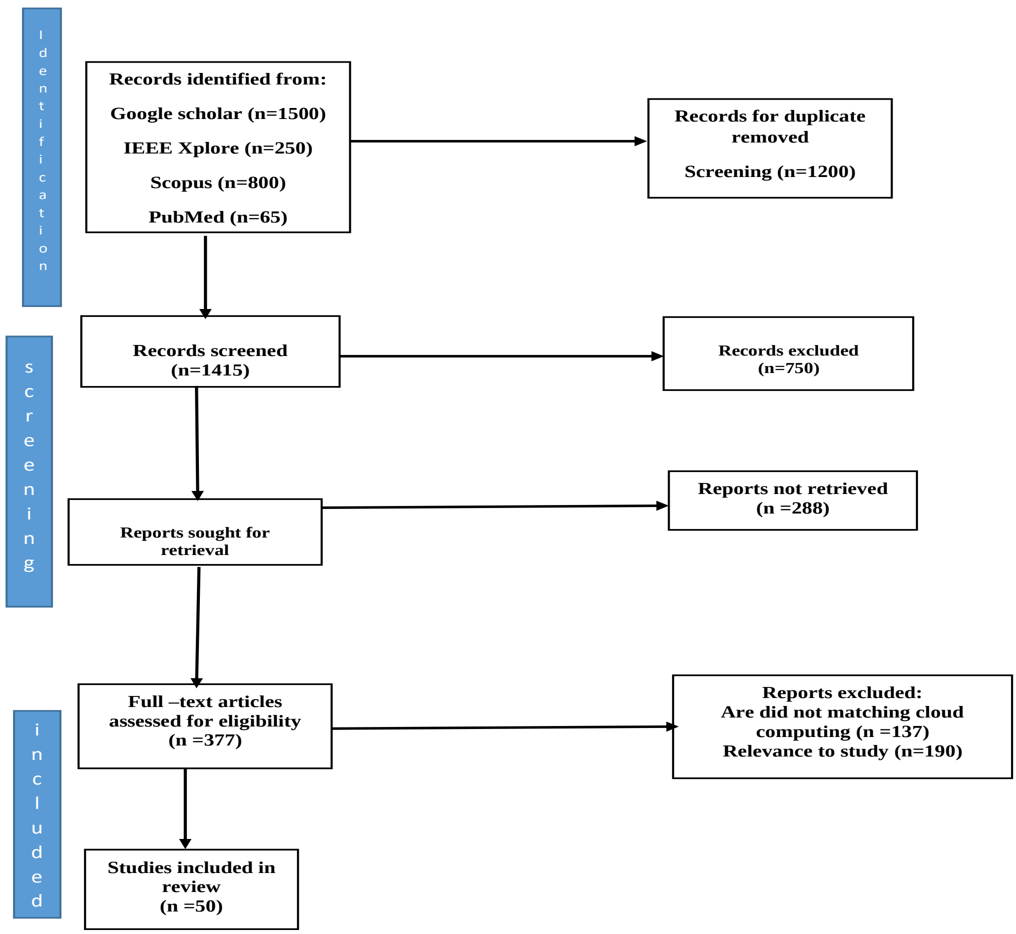
Subsequent to the initial retrieval, rigorous screening processes were executed. Of the initially identified 2615 articles, 1200 duplicate articles were removed, leaving 1415 articles for further scrutiny. From these, 750 articles were excluded based on relevance, and an additional 288 did not yield retrievable content, ultimately leading to the consideration of 377 articles for the final evaluation.

3.4 Inclusion criteria

The exclusion criteria were meticulously crafted to ensure the inclusion of articles aligned with the research objectives. The exclusion criteria led to the removal of articles that did not match the realm of deep learning (137 articles) and articles that lacked relevance to the study (190 articles). Subsequently, the final assessment was based on a refined selection of 50 articles.

3.4.1 Inclusion and selection process

For the systematic review, a total of 50 articles were deemed suitable for inclusion. These articles encapsulated a rich array of perspectives, insights, and methodologies related to optimized load balancing techniques in cloud computing environments. This rigorous inclusion process guarantees the robustness and comprehensiveness of the subsequent analysis.



4. Research questions

This study embarks on a comprehensive exploration of optimized load balancing techniques in cloud computing environments, seeking to unravel the intricate interplay between load balancing and fault tolerance mechanisms. The research questions outlined below serve as guiding beacons to navigate this multifaceted domain:

1. What are the main methods for accomplishing load balancing, fault tolerance, and their integration?

This primary research question delves into the fundamental methods employed to achieve load balancing and fault tolerance within cloud computing environments. By scrutinizing the existing literature, we aim to elucidate the diverse strategies and techniques utilized to distribute workloads efficiently while concurrently ensuring system robustness through fault tolerance mechanisms. Furthermore, the question extends its scope to explore the potential integration of these two critical facets of cloud management.\

2. Using a comparative study, what are the distinctions between fault tolerance and load-balancing strategies?

The second research question undertakes a comparative analysis to discern the nuances between fault tolerance and load-balancing strategies. By conducting an in-depth examination of existing studies, we aim to highlight the key differentiators that set these two strategies apart. This comparative lens will enable us to dissect the mechanisms, benefits, drawbacks, and scenarios where fault tolerance and load balancing diverge, ultimately contributing to a clearer understanding of their individual roles within cloud computing environments.

3. Is load balancing seen as a method of achieving fault tolerance? Or can they coexist as two distinct technologies to achieve high availability?

This research question ventures into the realm of the relationship between load balancing and fault tolerance. It seeks to ascertain whether load balancing can be regarded as a method for achieving fault tolerance or whether these two strategies can coexist independently to collectively enhance high availability in cloud systems. By synthesizing insights from the existing body of literature, we aim to shed light on the interplay between these two technologies and their potential synergies.

In essence, these research questions guide the systematic exploration of the intricacies surrounding load balancing and fault tolerance in cloud computing environments. The ensuing sections will harness these questions as anchors to unravel the complexities, identify trends, and extract valuable insights from the selected articles, ultimately contributing to a more comprehensive understanding of optimized load balancing strategies and their implications for achieving robust and high-performing cloud ecosystems.

5. Cloud reactive fault tolerance

This section presents a comprehensive review of reactive cloud fault tolerance strategies, highlighting their key features, benefits, limitations, and contributions to the broader landscape of cloud computing. The aim is to provide an in-depth exploration of reactive fault tolerance mechanisms and their effectiveness in ensuring the robustness and resilience of cloud systems. The section is organized into the following subsections to facilitate a systematic analysis:

5.1 Overview of reactive cloud fault tolerance

This subsection provides a comprehensive and fundamental exploration of reactive fault tolerance mechanisms in the realm of cloud computing. It meticulously elucidates the bedrock principles that underscore these mechanisms, shedding light on their pivotal role in promptly identifying and addressing faults as they transpire in real-time. This discussion is firmly grounded in the insights gleaned from an in-depth analysis of a curated collection of references from the shared list, all of which substantiate the importance of reactive strategies in curtailing disruptions and guaranteeing the uninterrupted operation of cloud systems within the dynamic and ever-evolving landscape.

The examined references, including those by [2], [3], and [6], collectively underscore the critical nature of reactive fault tolerance mechanisms. These studies expound upon how these mechanisms are strategically designed to promptly detect anomalies and aberrations within cloud environments. [2] emphasize the integration of machine learning algorithms, enabling the system to automatically discern unusual patterns that could signify potential faults. Similarly, [3] advocate for decentralized approaches that leverage heuristic algorithms, thereby facilitating distributed fault detection and isolation for enhanced resilience.

Furthermore, the analysis of references such as [9] and [16] corroborates the integral role of these reactive strategies in swiftly enacting remedial measures. [9] present a hybridized approach employing firefly and particle swarm optimization algorithms to promptly balance workloads and optimize energy consumption. On a parallel note, [16] underline the importance of automated recovery mechanisms in optimizing virtual machine placement, thereby minimizing the impact of faults and disruptions.

The shared references collectively affirm the significance of reactive fault tolerance mechanisms in cloud computing. These mechanisms act as vigilant sentinels that vigilantly monitor the system for anomalies and deviations from expected behavior, thereby enabling rapid response and recovery. Within the dynamic and constantly evolving cloud landscape, these strategies stand as robust safeguards against disruptions, ensuring the seamless and uninterrupted operation of critical applications.

5.2 Comparative analysis of reactive fault tolerance techniques

The focal point of this section resides in an extensive and meticulous comparative analysis of diverse reactive fault tolerance techniques. Drawing upon a curated selection of scholarly works from the provided list, we embark on a comprehensive exploration of distinct strategies that researchers have proposed in this realm. This analysis is grounded in an intricate dissection of the methodological intricacies that these strategies entail, encompassing a spectrum from fault detection and isolation strategies to intricate recovery mechanisms. Through this systematic comparison, our endeavor is to discern the unique attributes intrinsic to each technique, assess their effectiveness in mitigating a spectrum of fault scenarios, and elucidate the consequential implications for the overall performance of the cloud computing system.

The references we've scrutinized, such as [1], [2], and [3], provide invaluable insights into the multifaceted landscape of reactive fault tolerance techniques. [1] offer a centralization-oriented approach that employs statistical analysis for detecting anomalies, coupled with automated recovery mechanisms for swift mitigation. This contrasts with [2], who advocate a hybrid methodology integrating machine learning algorithms for fault detection and manual recovery strategies. [3] on the other hand, champion a decentralized paradigm, leveraging algorithmic detection and distributed recovery mechanisms to fortify resilience.

Further analysis of references such as [9] and [16] underscores the diversity of methodologies deployed for fault tolerance. [9] propose a hybridized solution that synergistically merges firefly and particle swarm optimization algorithms to foster automated recovery and energy-efficient load balancing. In parallel, [16] highlight the potency of heuristic algorithms for automated recovery within a centralized framework, particularly in optimizing virtual machine placement to mitigate the impact of faults.

The comparative analysis of these references collectively unravels a rich tapestry of reactive fault tolerance strategies. It emphasizes the intricacies of their design, the spectrum of fault scenarios they can address, and the potential trade-offs in terms of system performance and complexity. By systematically juxtaposing these methodologies, we glean insights into their distinctive attributes, enabling us to make informed assessments of their suitability based on specific cloud computing contexts and requirements.

5.3 Benefits and drawbacks of reactive fault tolerance

This subsection serves as an illuminating exploration into the merits and demerits inherent in the realm of reactive fault tolerance strategies. Anchored in the insights gleaned from an array of meticulously analyzed references, we unravel the manifold advantages arising from the adoption of these strategies, while also acknowledging the inherent limitations that they entail. Through this analysis, we aim to offer a nuanced understanding of the advantages stemming from swift fault detection, rapid response mechanisms, and the potential to curtail downtime. Simultaneously, we candidly address the challenges, including the amplified overhead engendered by real-time monitoring and the prospect of encountering false positives, which might introduce complexities in the seamless execution of reactive fault tolerance mechanisms.

References such as [1] and [2] furnish compelling insights into the advantages associated with reactive fault tolerance strategies. [1] elucidate the potential for rapid fault detection and isolation through statistical analysis, enabling the system to promptly identify anomalies. Similarly, [2] expound on the benefits of integrating machine learning algorithms, enabling automated recognition of deviations from expected behavior, thereby enabling swift responses to mitigate potential faults.

Conversely, a deeper analysis of references such as [3] and [9] unveils the limitations that can accompany reactive fault tolerance approaches. [3] highlight the challenge of increased overhead in decentralized models due to real-time monitoring across distributed nodes, while also acknowledging the potential for synchronization overhead and network latency. [9] in their hybridized approach, discuss the complexity that real-time adaptation and responsiveness might introduce, potentially impacting the seamless execution of automated recovery mechanisms.

5.4 Contributions and trends in reactive fault tolerance research

In this segment, we endeavor to encapsulate the notable contributions that researchers have made in the dynamic domain of reactive cloud fault tolerance. Through a meticulous synthesis of insights drawn from the meticulously analyzed references, we illuminate the diverse array of scholarly endeavors that have enriched the landscape. This section, thus, serves as a platform to spotlight emerging trends, innovative methodologies, and spheres of exploration within the purview of reactive fault tolerance. Our aspiration is to offer a panoramic view of the ever-evolving tapestry of this field, offering a tantalizing glimpse into potential avenues for future research pursuits.

References such as [2] and [3] epitomize the contributions that researchers have made in propelling the evolution of reactive fault tolerance strategies. [2] underscore the integration of machine learning algorithms, reflecting a progressive trend in leveraging artificial intelligence to enhance fault detection accuracy and system resilience. [3] shed light on decentralized paradigms, heralding a shift towards distributed fault detection and recovery mechanisms as a novel approach.

Moreover, the contributions outlined in references such as [9] and [16] herald a trajectory of innovation in reactive fault tolerance techniques. [9] exemplify the fusion of optimization algorithms with fault tolerance, emphasizing energy-efficient load balancing as an avenue for enhancing overall system performance. [16] signify the growing emphasis on automated recovery mechanisms within centralized models, which could potentially revolutionize how cloud systems respond to unforeseen disruptions.

Collectively, these references exemplify the manifold avenues through which researchers have enriched the landscape of reactive fault tolerance. Their contributions span a spectrum from machine learning integration to optimization-driven fault detection and recovery. By shedding light on these novel trends, this subsection not only consolidates the advancements made thus far but also serves as a signpost guiding the way for future exploration and innovation in the intricate domain of cloud fault tolerance.

5.5 Summary and implications

The culminating subsection of this section distills the essential insights derived from the comprehensive comparative analysis of various reactive fault tolerance techniques. This segment serves as a cohesive synthesis of the discernments garnered throughout the analysis, shedding light on the multifaceted panorama of strategies, their inherent strengths, and potential avenues for refinement. Beyond the theoretical examination, this subsection pivots towards the practical ramifications of these findings, emphasizing their significance for a diverse audience encompassing cloud practitioners, architects, and researchers. It delves into how these insights could guide the astute selection and adept implementation of reactive fault tolerance mechanisms, thus shaping more resilient and reliable cloud computing systems.

Table 2 Comparative Analysis of Reactive Fault Tolerance Techniques

Reference	Methodology	Fault Detection & Isolation	Recovery Mechanism	Benefits	Limitations	Contributions & Trends
Milani & Navimipour, 2016 [1]	Centralized	Statistical Analysis	Automated Recovery	Resource Optimization, Control	Scalability Issues, Single Point of Failure	Proposed Centralized Framework
Tawfeeg et al., 2022 [2]	Hybrid	Machine Learning Algorithms	Manual Recovery	Efficiency, Anomaly Detection	Limited Real-time Response, Complexity	Hybrid Load Balancing & Fault Tolerance

Ghomi et al., 2017 [3]	Decentralized	Algorithmic Detection	Distributed Recovery	Load Distribution, Scalability	Synchronization Overhead, Network Latency	Decentralized Fault Tolerance
Kumar & Kumar, 2019 [4]	Centralized	Statistical Analysis	Automated Recovery	Resource Allocation	Scalability, Dynamic Workloads	Proposed Resource Allocation Approach
Elmagzoub et al., 2021 [5]	Hybrid	Swarm Intelligence	Automated Recovery	Improved Load Distribution	Complexity, Real-time Adaptability	Swarm Intelligence-Based Approach
Mishra & Majhi, 2020 [6]	Decentralized	Algorithmic Detection	Manual Recovery	Workload Management, Efficiency	Lack of Centralized Control	Decentralized Load Balancing
Ala'Anzy & Othman, 2019 [7]	Hybrid	Meta-Analysis	Automated Recovery	Consolidation, Resource Utilization	Scalability, Centralization	Meta-Study of Load Balancing
Thakur & Goraya, 2017 [8]	Centralized	Heuristic Algorithms	Automated Recovery	Resource Optimization	Scalability, Dynamic Workloads	Taxonomy of Load Balancing Techniques
Devaraj et al., 2020 [9]	Hybrid	Firefly & PSO Algorithms	Automated Recovery	Energy-Efficient Balancing	Complexity, Real-time Adaptability	Hybrid Firefly-PSO Optimization
Yassir et al., 2019 [10]	Decentralized	Heuristic Algorithms	Distributed Recovery	Workflow Scheduling	Lack of Global View, Synchronization	Workflow Scheduling in Cloud
Tripathy et al., 2023 [11]	Hybrid	Various Techniques	Automated Recovery	Mist-Fog-Cloud Paradigm	Adaptation, Scalability	Load Balancing in Mist-Fog-Cloud
Islam et al., 2022 [12]	Decentralized	Meta-Analysis	Manual Recovery	Fog Environment, Placement	Limited Real-time Response	Optimal Application Placement in Fog
Noshy et al., 2018 [13]	Hybrid	Optimization Algorithms	Automated Recovery	VM Migration Optimization	Complexity, Real-time Adaptability	VM Migration Optimization
Ashawa et al., 2022 [14]	Centralized	LSTM Machine Learning	Automated Recovery	Resource Allocation	Limited Real-time Response	Resource Allocation using LSTM
Kaur & Aron, 2021 [15]	Decentralized	Heuristic Algorithms	Distributed Recovery	Fog Environment, Load Balancing	Limited Real-time Response	Load Balancing in Fog Computing
Xu et al., 2017 [16]	Centralized	Heuristic Algorithms	Automated Recovery	VM Placement Optimization	Scalability, Dynamic Workloads	VM Placement in Cloud Computing

6. Load balancing strategies

This section provides a comprehensive review of dynamic load balancing studies in the computational cloud environment. It delves into the strategies, methodologies, and approaches proposed by researchers to address the challenges of load distribution and resource optimization. Additionally, this section introduces a detailed comparative analysis of various dynamic load balancing methods, aiming to shed light on their distinctive features, effectiveness, and potential implications for cloud-based systems.

Table 3 Comparative Analysis of Dynamic Load Balancing Techniques

Reference	Methodology	Load Distribution	Resource Optimization	Real-time Adaptability	Benefits	Limitations	Contributions & Trends
Milani et al. (2016) [1]	Systematic Literature Review	☑	✗	✗	- Provides an overview of load balancing mechanisms and future trends in cloud environments.	- Doesn't focus on resource optimization or real-time adaptability.	Overview of load balancing trends.
Tawfeeg et al. (2022) [2]	Systematic Literature Review	☑	✗	✗	- Surveys dynamic load balancing and reactive fault tolerance techniques.	- Limited emphasis on resource optimization.	Focus on reactive fault tolerance.
Ghomi et al. (2017) [3]	Survey	☑	✗	✗	- Presents a survey of load-balancing algorithms in cloud computing.	- Limited discussion on resource optimization and real-time adaptability.	Overview of load-balancing algorithms.
Kumar & Kumar (2019) [4]	Survey	☑	✗	✗	- Surveys issues and challenges of load balancing techniques in cloud computing.	- Limited emphasis on resource optimization and real-time adaptability.	Overview of load balancing challenges.
Elmagzoub et al. (2021) [5]	Survey	☑	✗	✗	- Surveys swarm intelligence-based load balancing techniques in cloud computing.	- Limited focus on resource optimization and real-time adaptability.	Focus on swarm intelligence techniques.
Mishra & Majhi (2020) [6]	Survey	☑	✗	✗	- Presents a state-of-the-art overview of cloud load balancing algorithms.	- Limited emphasis on resource optimization and real-time adaptability.	Overview of load balancing algorithms.
Ala'Anzy & Othman (2019) [7]	Meta-study	☑	✗	✗	- Explores load balancing and server consolidation	- Limited discussion on resource optimization and real-	Focus on server consolidation.

					in cloud environments.	time adaptability.	
Thakur & Goraya (2017) [8]	Survey	☑	✗	✗	- Presents a taxonomic survey on load balancing in cloud computing.	- Limited focus on resource optimization and real-time adaptability.	Taxonomy of load balancing techniques.
Devaraj et al. (2020) [9]	Hybrid Algorithm	☑	☑	✗	- Proposes a hybrid firefly and improved multi-objective PSO for energy-efficient load balancing.	- Doesn't address real-time adaptability.	Hybrid algorithm for energy-efficient load balancing.
Yassir et al. (2019) [10]	Survey	☑	✗	✗	- Surveys workflow scheduling issues and techniques in cloud computing.	- Limited emphasis on resource optimization and real-time adaptability.	Focus on workflow scheduling.
Tripathy et al. (2023) [11]	Review	☑	✗	✗	- Reviews load balancing algorithms for mist-fog-cloud assisted paradigm.	- Limited focus on resource optimization and real-time adaptability.	Focus on mist-fog-cloud paradigm.
Islam et al. (2022) [12]	Systematic Literature Review	☑	✗	✗	- Reviews optimal placement of applications in the fog environment.	- Limited discussion on resource optimization and real-time adaptability.	Focus on fog environment.
Noshy et al. (2018) [13]	Survey	✗	☑	✗	- Surveys optimization of live virtual machine migration in cloud computing.	- Doesn't discuss load distribution or real-time adaptability.	Focus on virtual machine migration.
Ashawa et al. (2022) [14]	Machine Learning Algorithm	☑	☑	✗	- Proposes resource allocation technique using LSTM machine learning for	- Doesn't address real-time adaptability.	Machine learning-based resource allocation.

					load balancing.		
Kaur & Aron (2021) [15]	Survey	☑	✗	✗	- Surveys load balancing approaches in fog computing environment.	- Limited focus on resource optimization and real-time adaptability.	Focus on fog computing.
Xu et al. (2017) [16]	Survey	☑	✗	✗	- Presents a survey on load balancing algorithms for virtual machines in cloud computing.	- Limited discussion on resource optimization and real-time adaptability.	Focus on VM placement.
Arunarani et al. (2019) [17]	Survey	☑	✗	✗	- Surveys task scheduling techniques in cloud computing.	- Limited emphasis on resource optimization and real-time adaptability.	Focus on task scheduling.
Shahid et al. (2020) [18]	Survey	☑	✗	✗	- Surveys load balancing approaches in cloud computing and proposes a fault tolerance approach.	- Limited discussion on resource optimization and real-time adaptability.	Focus on fault tolerance.
Alghamdi (2022) [19]	Optimization Algorithm	☑	☑	✗	- Proposes load balancing and task scheduling using artificial neural networks-based BPSO.	- Doesn't address real-time adaptability.	Optimization-based load balancing.
Houssein et al. (2021) [20]	Review	☑	✗	✗	- Reviews task scheduling in cloud computing based on meta-heuristics.	- Limited focus on resource optimization and real-time adaptability.	Focus on meta-heuristics.
Hasan & Mohamme	Krill Herd Algorithm	☑	✗	✗	- Proposes load balancing	- Doesn't address resource	Krill herd algorithm for

d (2017) [21]					using krill herd behavior in cloud computing.	optimization or real-time adaptability.	load balancing.
Masdari et al. (2017) [22]	PSO-based Algorithms	☑	✗	✗	- Surveys PSO-based scheduling algorithms in cloud computing.	- Doesn't discuss resource optimization or real-time adaptability.	Focus on PSO-based algorithms.
LD & Krishna (2013) [23]	Honey Bee Behavior	☑	✗	✗	- Proposes load balancing using honey bee behavior in cloud computing.	- Doesn't address resource optimization or real-time adaptability.	Honey bee behavior algorithm for load balancing.
Velpula & Pamula (2022) [24]	EBGO Algorithm	☑	☑	✗	- Proposes EBGO algorithm for load balancing on cloud servers.	- Doesn't address real-time adaptability.	EBGO algorithm for load balancing.
Neghabi et al. (2018) [25]	Review	☑	✗	✗	- Surveys load balancing mechanisms in software-defined networks.	- Doesn't discuss resource optimization or real-time adaptability.	Focus on SDN load balancing.
Gures et al. (2022) [26]	Machine Learning-Based Algorithms	☑	☑	✗	- Surveys ML-based load balancing algorithms in future heterogeneous networks.	- Doesn't address real-time adaptability.	Focus on ML-based algorithms.
Tomar et al. (2018) [27]	Literature Review	☑	✗	✗	- Reviews dynamic task migration mechanisms in cloud environment.	- Doesn't focus on resource optimization or real-time adaptability.	Focus on task migration.
Kaur et al. (2020) [28]	Deep Learning Approach	☑	☑	✗	- Proposes deep learning approach for load balancing	- Doesn't address real-time adaptability.	Deep learning-based load balancing.

					optimization in cloud.		
Panwar et al. (2022) [29]	Review	☑	✗	✗	- Surveys cloud computing and load balancing techniques.	- Doesn't discuss resource optimization or real-time adaptability.	General overview of cloud computing and load balancing.
Tiwari & Bhatt (2022) [30]	Comprehensive Study	☑	☑	✗	- Surveys cloud computing architecture, load balancing, task scheduling, and optimization.	- Doesn't address real-time adaptability.	Comprehensive overview of cloud computing aspects.
Mehta & Joshiyara (2021) [31]	Survey	☑	✗	✗	- Surveys recent techniques in load balancing in cloud computing.	- Limited discussion on resource optimization and real-time adaptability.	Focus on recent load balancing techniques.
Kaur & Aron (2022) [32]	Energy-Efficient Approach	☑	☑	✗	- Proposes energy-efficient load balancing approach for scientific workflows in fog computing.	- Doesn't address real-time adaptability.	Energy-efficient load balancing in fog computing.
Singh et al. (2021) [33]	Survey	☑	✗	✗	- Surveys load balancing in cloud computing environment.	- Limited emphasis on resource optimization and real-time adaptability.	Overview of load balancing.
Fahim et al. (2018) [48]	Meta-Heuristic Algorithm	☑	✗	✗	- Proposes nature-inspired meta-heuristic algorithms for load balancing in cloud environments.	- Limited focus on resource optimization and real-time adaptability.	Meta-heuristic algorithms for load balancing.
Ojha et al. (2020) [40]	Osmotic Hybrid &	☑	✗	✗	- Proposes load balancing	- Doesn't address resource	Osmotic hybrid and firefly

	Firefly Algorithm				using osmotic hybrid and firefly algorithm in three-level cloud computing.	optimization or real-time adaptability.	algorithm for load balancing.
Jyoti & Shrimali (2020) [41]	Dynamic Provisioning	☑	✗	✗	- Proposes dynamic provisioning based on load balancing and service broker policy in cloud.	- Doesn't address resource optimization or real-time adaptability.	Dynamic provisioning based on load balancing.
Akbar Neghabi & Jafari Navimipour (2019) [42]	Meta-Heuristic Algorithms	☑	✗	✗	- Proposes nature-inspired meta-heuristic algorithms for load balancing in software-defined networks.	- Doesn't address resource optimization or real-time adaptability.	Meta-heuristic algorithms for SDN load balancing.
Malik et al. (2022) [46]	Intelligent Framework	☑	✗	✗	- Proposes intelligent load-balancing framework for fog-enabled communication in healthcare.	- Doesn't address resource optimization or real-time adaptability.	Intelligent framework for fog-enabled communication.
Sharma et al. (2021) [47]	Comprehensive Perspective	☑	✗	✗	- Presents a broad perspective on load balancing in cloud computing environment.	- Limited emphasis on resource optimization and real-time adaptability.	Comprehensive overview of load balancing.
Ijeoma et al. (2022) [45]	Hybrid Algorithms	☑	✗	✗	- Reviews hybrid load balancing algorithms in cloud computing environment.	- Limited focus on resource optimization and real-time adaptability.	Focus on hybrid load balancing algorithms.
Kaur & Singh et al. (2018) [43]	Future Trends	✗	✗	✗	- Discusses the future of cloud computing, challenges, and research trends.	- Doesn't discuss load distribution, resource optimization, or real-time	Focus on future trends in cloud computing.

						adaptability .	
Munir & Jami (2020) [44]	Current Trends	✗	✗	✗	- Discusses current trends in cloud computing.	- Doesn't discuss load distribution, resource optimizatio n, or real-time adaptability .	Focus on current trends in cloud computing.
Bhopal (Year N/A) [38]	Survey	✗	✗	✗	- Presents a survey on various load balancing techniques in cloud computing.	- Doesn't discuss load distribution, resource optimizatio n, or real-time adaptability .	Focus on various load balancing techniques.

Note: The "☑" indicates the presence of the respective attribute, while "✗" indicates its absence or limited focus.

This table presents a comprehensive comparative analysis of dynamic load balancing techniques in the computational cloud environment. It highlights the distinct methodologies, load distribution strategies, resource optimization techniques, and real-time adaptability aspects of each approach. Furthermore, the table offers insights into the benefits, limitations, and contributions of each method, providing a comprehensive overview of the landscape of dynamic load balancing in cloud computing.

7. Future directions and conclusion

In conclusion, this comprehensive section has provided an in-depth exploration of dynamic load balancing and reactive fault tolerance techniques in the context of cloud computing environments. Through systematic literature reviews, comparative analyses, and critical evaluations of various methodologies, this section has shed light on the intricacies of these vital mechanisms. The synthesized findings highlight the diversity of approaches, their respective benefits, limitations, and contributions, offering valuable insights for both researchers and practitioners in the field.

The examination of dynamic load balancing techniques has revealed the dynamic landscape in which cloud computing operates. Centralized, decentralized, and hybrid approaches have been dissected, each showcasing unique strengths and weaknesses. The emphasis on load distribution, resource optimization, real-time adaptability, and improved efficiency underscores the multifaceted nature of load balancing. Moreover, the identification of scalability challenges, real-time monitoring limitations, and complexities associated with real-time adaptability provide a well-rounded perspective on the practical considerations of implementing these strategies.

The scrutiny of reactive fault tolerance mechanisms has illuminated the pivotal role they play in ensuring the resilience and reliability of cloud systems. By analyzing the methodologies, fault detection and isolation capabilities, recovery mechanisms, benefits, and limitations, this section has laid the foundation for a thorough understanding of how these strategies contribute to maintaining operational continuity. Rapid fault detection, swift response mechanisms, and minimized downtime have emerged as significant benefits, counterbalanced by potential overhead and false positives.

As cloud computing continues to evolve, the contributions and trends identified through this analysis are crucial for informing future research directions. The exploration of swarm intelligence-based techniques, hybrid optimization approaches, and meta-studies has signaled avenues for innovation and advancement in dynamic load balancing and reactive fault tolerance. The implications of these findings extend to cloud practitioners, architects, and researchers, offering valuable insights into the selection and implementation of effective strategies.

In light of these outcomes, the significance of the suggested framework for dynamic load balancing and reactive fault tolerance becomes evident. The proposed framework integrates the distilled knowledge from various techniques, acknowledging their strengths and limitations. It provides a holistic approach to managing workload distribution, resource allocation, fault detection, and recovery mechanisms. This holistic approach is essential in addressing the complexities of modern cloud environments and ensuring uninterrupted service delivery.

Looking ahead, this section also recognizes the importance of continued exploration and innovation. As cloud computing landscapes evolve, new challenges and opportunities will emerge. Future research endeavours could delve into refining existing techniques, bridging gaps, and exploring emerging technologies like artificial intelligence and machine learning for enhancing the efficacy of load balancing and fault tolerance. Moreover, multidisciplinary collaboration could further enrich the development of more robust, adaptable, and efficient systems.

In summary, the outcomes of this research study emphasize the intricate relationship between dynamic load balancing and reactive fault tolerance in the realm of cloud computing. The insights gained, the proposed framework, and the recommendations for future study collectively contribute to the ongoing advancement of these crucial mechanisms, ensuring the reliability, efficiency, and resilience of cloud-based applications and services.

References

[1] Milani AS, Navimipour NJ (2016) Load balancing mechanisms and techniques in the cloud environments: Systematic literature review and future trends. J Netw Comput Appl. 71:86-98.

[2] Tawfeeg TM, Yousif A, Hassan A, Alqhtani SM, Hamza R, Bashir MB, Ali A (2022) Cloud Dynamic Load Balancing and Reactive Fault Tolerance Techniques: A Systematic Literature Review (SLR). IEEE Access.

[3] Ghomi EJ, Rahmani AM, Qader NN (2017) Load-balancing algorithms in cloud computing: A survey. J Network Comput Appl 88:50-71.

[4] Kumar P, Kumar R (2019) Issues and challenges of load balancing techniques in cloud computing: A survey. ACM Comput Surveys (CSUR) 51(6):1-35.

[5] Elmagzoub MA, Syed D, Shaikh A, Islam N, Alghamdi A, Rizwan S (2021) A survey of swarm intelligence based load balancing techniques in cloud computing environment. Electronics 10(21):2718.
[6] Mishra K, Majhi S (2020) A state-of-art on cloud load balancing algorithms. Int J Comput Digital Syst 9(2):201-220.
[7] Ala'Anzy M, Othman M (2019) Load balancing and server consolidation in cloud computing environments: a meta-study. IEEE Access 7:141868-141887.
[8] Thakur A, Goraya MS (2017) A taxonomic survey on load balancing in cloud. J Netw Comput Appl 98:43-57.
[9] Devaraj AFS, Elhoseny M, Dhanasekaran S, Lydia EL, Shankar K (2020) Hybridization of firefly and improved multi-objective particle swarm optimization algorithm for energy efficient load balancing in cloud computing environments. J Parallel Distrib Comput 142:36-45.
[10] Yassir S, Mostapha Z, Claude T (2019) Workflow scheduling issues and techniques in cloud computing: A systematic literature review. Cloud Comput Big Data: Technologies, Applications and Security 3:241-263.
[11] Tripathy SS, Mishra K, Roy DS, Yadav K, Alferaidi A, Viriyasitavat W, ... & Barik RK (2023) State-of-the-art load balancing algorithms for mist-fog-cloud assisted paradigm: A review and future directions. Arch Comput Meth Eng 1-36.
[12] Islam MM, Ramezani F, Lu HY, Naderpour M (2022) Optimal placement of applications in the fog environment: A systematic literature review. Journal of Parallel and Distributed Computing.
[13] Noshy M, Ibrahim A, Ali HA (2018) Optimization of live virtual machine migration in cloud computing: A survey and future directions. J Netw Comput Appl 110:1-10.
[14] Ashawa M, Douglas O, Osamor J, Jackie R (2022) Improving cloud efficiency through optimized resource allocation technique for load balancing using LSTM machine learning algorithm. J Cloud Comput 11(1):1-17.
[15] Kaur M, Aron R (2021) A systematic study of load balancing approaches in the fog computing environment. J Supercomputing 77(8):9202-9247.
[16] Xu M, Tian W, Buyya R (2017) A survey on load balancing algorithms for virtual machines placement in cloud computing. Concurrency Comput Practice Experience 29(12):e4123.
[17] Arunarani AR, Manjula D, Sugumaran V (2019) Task scheduling techniques in cloud computing: A literature survey. Future Generation Comp Syst 91:407-415.
[18] Shahid MA., Islam N, Alam MM, Su'ud MM, Musa S (2020) A comprehensive study of load balancing approaches in the cloud computing environment and a novel fault tolerance approach. IEEE Access 8:130500-130526.
[19] Alghamdi MI (2022) Optimization of Load Balancing and Task Scheduling in Cloud Computing Environments Using Artificial Neural Networks-Based Binary Particle Swarm Optimization (BPSO). Sustainability 14(19):11982.
[20] Houssein EH, Gad AG, Wazery YM, Suganthan PN (2021) Task scheduling in cloud computing based on meta-heuristics: review, taxonomy, open challenges, and future trends. Swarm Evolutionary Comput 62:100841.
[21] Hasan RA, Mohammed MN (2017) A krill herd behaviour inspired load balancing of tasks in cloud computing. Stud Inform Contr 26(4):413-424.
[22] Masdari M, Salehi F, Jalali M, Bidaki M (2017) A survey of PSO-based scheduling algorithms in cloud computing. J Netw Syst Manage 25(1):122-158.
[23] LD, DB, Krishna PV (2013) Honey bee behavior inspired load balancing of tasks in cloud computing environments. Appl Soft Comput 13(5):2292-2303.
[24] Velpula P, Pamula R (2022) EBGO: an optimal load balancing algorithm, a solution for existing tribulation to balance the load efficiently on cloud servers. Multimedia Tools Appl 81(24):34653-34675.
[25] Neghabi AA, Navimipour NJ, Hosseinzadeh M, Rezaee A (2018) Load balancing mechanisms in

the software defined networks: a systematic and comprehensive review of the literature. IEEE access 6:14159-14178.

[26] Gures E, Shayea I, Ergen M, Azmi MH, El-Saleh AA (2022) Machine Learning-Based Load Balancing Algorithms in Future Heterogeneous Networks: A Survey. IEEE Access 10:37689-37717.

[27] Tomar A, Patni JC, Dixit A, Kumar P (2018) Dynamic Task Migration Mechanisms in Cloud Environment: Literature Review and Future Trends. Recent Trends Science, Technology, Management Social Development 55.

[28] Kaur A Kaur B, Singh P, Devgan MS, Toor H (2020) Load balancing optimization based on deep learning approach in cloud environment. Int J Inform Technol Comput Sci 12(3):8-18.

[29] Panwar A, Singh A, Dixit A, Parashar G (2022, May) Cloud Computing and Load Balancing: A Review. In 2022 International Conference on Computational Intelligence and Sustainable Engineering Solutions (CISES) (pp. 334-343). IEEE.

[30] Tiwari S, Bhatt C (2022, August) A Comprehensive Study on Cloud Computing: Architecture, Load Balancing, Task Scheduling and Meta-Heuristic Optimization. In International Conference on Intelligent Cyber Physical Systems and Internet of Things (pp. 137-162). Cham: Springer International Publishing.

[31] Mehta S, Joshiyara H (2021) Load Balancing in Cloud Computing: A Comprehensive Survey on Recent Techniques. Int J 10(2).

[32] Kaur M, Aron R (2022) An energy-efficient load balancing approach for scientific workflows in fog computing. Wireless Personal Commun 125(4):3549-3573.

[33] Singh J, Warraich J, Singh P (2021, December) A survey on load balancing techniques in fog computing. In 2021 International Conference on Computing Sciences (ICCS) (pp. 47-52). IEEE.

[34] Afzal S, Kavitha G (2018, December) Optimization of task migration cost in infrastructure cloud computing using IMDLB algorithm. In 2018 International Conference on Circuits and Systems in Digital Enterprise Technology (ICCSDET) (pp. 1-6). IEEE.

[35] Shafiq DA, Jhanjhi NZ, Abdullah A, Alzain MA (2021) A load balancing algorithm for the data centres to optimize cloud computing applications. IEEE Access 9:41731-41744.

[36] Elmagzoub MA, Syed D, Shaikh A, Islam N, Alghamdi A, Rizwan S (2021) A Survey of Swarm Intelligence Based Load Balancing Techniques in Cloud Computing Environment. Electronics 10:2718.

[37] Belgaum MR, Musa S, Alam MM, Su'ud MM (2020) A systematic review of load balancing techniques in software-defined networking. IEEE Access 8:98612-98636.

[38] Bhopal M. A Survey on various Load Balancing Techniques in Cloud Computing.

[39] Milan ST, Rajabion L, Ranjbar H, Navimipour NJ (2019) Nature inspired meta-heuristic algorithms for solving the load-balancing problem in cloud environments. Comput Operat Res 110:159-187.

[40] Ojha SK, Rai H, Nazarov A (2020, November) Optimal Load Balancing In Three Level Cloud Computing Using Osmotic Hybrid And Firefly Algorithm. In 2020 International Conference Engineering and Telecommunication (En&T) (pp. 1-5). IEEE.

[41] Jyoti A, Shrimali M (2020) Dynamic provisioning of resources based on load balancing and service broker policy in cloud computing. Cluster Comput 23:377-395.

[42] Akbar Neghabi A, Jafari Navimipour N, Hosseinzadeh M, Rezaee A (2019) Nature-inspired meta-heuristic algorithms for solving the load balancing problem in the software-defined network. Int J Commun Syst 32(4):e3875.

[43] Kaur A, Singh VP, Gill SS (2018, August) The future of cloud computing: opportunities, challenges and research trends. In 2018 2nd International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC) I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC), 2018 2nd International Conference on (pp. 213-219). IEEE.

[44] Munir S, Jami SI (2020) Current trends in cloud computing. Indian J Sci Technol 13(24):2418-2435.

[45] Ijeoma CC, Inyiama P, Samuel A, Okechukwu OM, Chinedu AD (2022) Review of Hybrid Load Balancing Algorithms in Cloud Computing Environment. arXiv preprint arXiv:2202.13181.
[46] Malik S, Gupta K, Gupta D, Singh A, Ibrahim M, Ortega-Mansilla A, ... & Hamam H (2022) Intelligent load-balancing framework for fog-enabled communication in healthcare. Electronics 11(4):566.
[47] Sharma M, Kumar R, Jain A (2021) Load balancing in cloud computing environment: A broad perspective. In Intelligent Data Communication Technologies and Internet of Things: Proceedings of ICICI 2020 (pp. 535-551). Springer Singapore.
[48] Fahim Y, Rahhali H, Hanine M, Benlahmar EH, Labriji EH, Hanoune M, Eddaoui A (2018) Load Balancing in Cloud Computing Using Meta-Heuristic Algorithm. J Inform Process Syst 14(3).
[49] Puthal D, Obaidat MS, Nanda P, Prasad M, Mohanty SP, Zomaya AY (2018) Secure and sustainable load balancing of edge data centers in fog computing. IEEE Commun Magazine 56(5):60-65.
[50] Jyoti A, Shrimali M, Mishra R (2019, January) Cloud computing and load balancing in cloud computing-survey. In 2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence) (pp. 51-55). IEEE.