Indian Journal of Natural Sciences *www.tnsroindia.org.in ©IJONS*

Vol.15 / Issue 86 / Oct / 2024 *International Bimonthly (Print) – Open Access* *ISSN: 0976 – 0997*

RESEARCH ARTICLE

Enhancing Quality of Service in Cloud Computing: A Comprehensive Review of Techniques and Challenges

R. Jeya[1]* and V. Baby Deepa[2]

[1]Research Scholar [Part Time], PG and Research Department of Computer Science, Government Arts College (Autonomous), Karur (Affiliated to Bharathidasan University, Tiruchirappalli), Tamil Nadu, India

[2]Research Advisor& Assistant Professor, PG and Research Department of Computer Science, Government Arts College (Autonomous), Karur (Affiliated to Bharathidasan University, Tiruchirappalli), Tamil Nadu, India.

Received: 21 Aug 2024 Revised: 07 Jul 2024 Accepted: 26 Oct 2024

***Address for Correspondence**
R. Jeya
Research Scholar [Part Time],
PG and Research Department of Computer Science,
Government Arts College (Autonomous), Karur
(Affiliated to Bharathidasan University, Tiruchirappalli),
Tamil Nadu, India
E.Mail: jeyarmca@gmail.com

ABSTRACT

The rapid growth of cloud computing has revolutionized data storage, application hosting, and resource management, making it a cornerstone of modern digital infrastructure. However, the adoption of cloud services introduces critical challenges related to performance optimization, Quality of Service (QoS), and security. This literature review aims to explore various strategies that have been proposed to improve cloud computing performance while maintaining high QoS standards and robust security measures. The review delves into techniques such as dynamic resource allocation, virtualization, load balancing, and traffic management to enhance system efficiency and scalability. It also examines encryption protocols, authentication mechanisms, and intrusion detection systems aimed at fortifying cloud security. By analyzing current research, this review highlights the trade-offs between performance and security, offering insights into how emerging technologies such as machine learning, blockchain, and edge computing are being integrated to strike a balance. The findings provide a comprehensive understanding of the existing frameworks and future directions for optimizing cloud performance without compromising on QoS and security requirements.

Keywords: Cloud Security, Quality of Service, Security

INTRODUCTION

Cloud computing has become an integral part of modern information technology infrastructures, enabling organizations and individuals to access and manage data, applications, and services over the internet. Its promise of on-demand resource availability, scalability, cost efficiency, and flexibility has driven widespread adoption across various industries, including healthcare, finance, education, and entertainment. However, as reliance on cloud computing increases, so do the challenges related to ensuring high levels of performance, Quality of Service (QoS), and security. Performance in cloud computing refers to the ability of cloud systems to provide users with fast, efficient, and reliable access to resources while handling large-scale workloads. This aspect is crucial in cloud environments where unpredictable spikes in demand can strain resources. To mitigate these issues, cloud service providers (CSPs) deploy performance enhancement techniques such as dynamic resource allocation, load balancing, and virtualization, which aim to maximize resource utilization and maintain optimal system throughput. However, maintaining high performance often comes at the expense of other critical factors, such as QoS and security [1] [2].

Quality of Service (QoS) represents the overall service quality experienced by end-users, encompassing elements such as availability, latency, bandwidth, and error rates. It is essential in maintaining user satisfaction and ensuring seamless operation for businesses relying on cloud services. The dynamic and distributed nature of cloud environments complicates QoS management, requiring sophisticated algorithms for resource allocation, traffic management, and real-time monitoring. Ensuring consistent QoS across geographically dispersed data centers, varying network conditions, and diverse workloads remains a complex challenge for CSPs [3]. At the same time, security in cloud computing is paramount, given the sensitivity of the data and applications hosted in these environments. Organizations entrust CSPs with confidential data, making cloud infrastructures attractive targets for cyberattacks. Cloud security encompasses a wide range of concerns, including data protection, user authentication, network security, and compliance with privacy regulations. Techniques such as encryption, intrusion detection, access control, and multi-factor authentication are employed to safeguard cloud environments. However, achieving high levels of security often involves overhead that can negatively impact performance, creating a trade-off between security measures and the efficiency of cloud operations [4] [5].

In this context, the interplay between performance, QoS, and security becomes a critical area of concern for researchers and practitioners alike. Improving performance while maintaining a robust QoS and ensuring high security levels presents a unique set of challenges. For instance, adding security measures such as encryption can lead to increased computational load, thus reducing performance. Similarly, prioritizing performance optimization might lead to gaps in security or compromise on QoS aspects like latency or service availability. This literature review seeks to explore the existing body of research that addresses these interconnected challenges in cloud computing. By examining state-of-the-art methodologies and technologies, including machine learning-based optimization techniques, blockchain for decentralized security, and edge computing for distributed workloads, the review will highlight approaches aimed at enhancing cloud performance without compromising QoS or security. It will also identify key trends and future directions in the field, particularly the integration of emerging technologies to balance performance, QoS, and security requirements in cloud environments.

Background Study of Cloud Security

Cloud computing has become a cornerstone of modern digital infrastructure, enabling organizations to deploy, store, and process data with unprecedented scalability, flexibility, and cost efficiency. However, as cloud adoption grows, so does the complexity of ensuring the security of these vast, distributed systems. Cloud security encompasses a broad range of technologies, protocols, and policies designed to protect data, applications, and infrastructure from a wide array of threats. The unique characteristics of cloud environments—such as multi-tenancy, on-demand resource provisioning, and remote access—create significant challenges for securing data integrity, confidentiality, and availability.

The Evolution of Cloud Security

Cloud security evolved alongside the development of cloud computing technologies. Initially, concerns about security were among the biggest barriers to cloud adoption, with enterprises hesitant to entrust sensitive data to third-party service providers. Over time, cloud providers invested heavily in developing security mechanisms that address both traditional IT security concerns and the novel risks introduced by cloud infrastructures [6] [7]. Traditional data centers operated within a single organizational boundary, where physical and network access controls could be tightly managed. In contrast, cloud environments are shared, decentralized, and accessible over the internet, raising new challenges in protecting against insider threats, external attacks, and accidental data leaks. To mitigate these risks, cloud service providers (CSPs) have implemented a range of security measures, such as encryption, identity management, access controls, and auditing tools. However, securing cloud environments is not solely the responsibility of CSPs; it requires a collaborative effort between providers and users, each responsible for different layers of the security model.

Shared Responsibility Model

One of the foundational concepts in cloud security is the Shared Responsibility Model. In this model, cloud service providers and customers share the responsibility of securing data and infrastructure, though the scope of responsibility varies based on the cloud service model used—Infrastructure as a Service (IaaS), Platform as a Service (PaaS), or Software as a Service (SaaS). IaaS: In an IaaS model, the cloud provider manages the physical data center, servers, and network, while the customer is responsible for securing the operating system, applications, and data. This model provides the greatest control for the user but also demands the most comprehensive security management from the customer's side.

PaaS: In PaaS, the provider also manages the underlying infrastructure and platform (such as operating systems and middleware), leaving customers responsible mainly for securing their applications and data. Security responsibilities here are somewhat reduced for customers but still involve considerable attention to data protection and application security. SaaS: In a SaaS model, the cloud provider handles nearly everything—network, servers, applications, and data storage—while the customer is generally responsible for securing their access credentials and managing user permissions. SaaS users are at the mercy of the provider's security measures, making it critical for customers to carefully evaluate their provider's security practices.

Key Cloud Security Threats

As cloud adoption grows, so does the landscape of threats targeting cloud environments. Some of the most prominent security threats include:

Data Breaches: Data breaches remain one of the most significant threats to cloud security. Multi-tenant cloud environments, where data from multiple customers resides on shared infrastructure, increase the risk of unauthorized access. Breaches can occur due to poor configuration, weak access controls, or vulnerabilities in cloud services. High-profile data breaches, such as the Capital One incident in 2019, underscore the risks posed by misconfigurations in cloud infrastructure [8].

Data Loss: Data loss can occur for a variety of reasons, including accidental deletion, hardware failures, software bugs, and malicious attacks (such as ransomware). In a cloud context, ensuring data redundancy and backup procedures are critical for minimizing the risk of permanent data loss. While replication across multiple data centers can help mitigate this risk, it also requires careful management to ensure consistency and data integrity.

Insider Threats: Cloud environments are particularly vulnerable to insider threats due to the vast number of users with varying levels of access. Employees or administrators of cloud providers, as well as users within an organization, may intentionally or unintentionally misuse their access, leading to data breaches or service disruptions. Insider threats can be mitigated by strong access controls, logging, and continuous monitoring.

Denial of Service (DoS) Attacks: Distributed Denial of Service (DDoS) attacks, which aim to overwhelm cloud services with traffic, can render services unavailable to legitimate users. Cloud providers offer DDoS mitigation tools, such as traffic filtering and load balancing, but large-scale attacks can still cause significant disruptions, especially for smaller providers [9].

Insecure APIs: Application Programming Interfaces (APIs) are widely used in cloud environments to enable integration and automation. However, insecure APIs—those lacking proper authentication, encryption, or access controls—can expose cloud systems to cyberattacks. Ensuring the security of APIs is a critical aspect of cloud security, requiring regular auditing and testing for vulnerabilities.

Account Hijacking: Weak or stolen credentials remain a common attack vector in cloud security. Once attackers gain access to cloud accounts, they can manipulate data, steal sensitive information, or perform malicious activities under the guise of legitimate users. Multifactor authentication (MFA), encryption, and strict access controls are essential defenses against account hijacking.

Background Study of the Data Replication Methodologies

Cloud computing offers numerous advantages such as scalability, flexibility, and cost efficiency. However, ensuring **Quality of Service (QoS)** in a cloud computing environment remains one of the most significant challenges faced by both Cloud Service Providers (CSPs) and users. QoS in cloud computing refers to the overall performance of the cloud system, which includes metrics such as **availability**, **reliability**, **latency**, **bandwidth**, and **error rates**. A high QoS guarantees that users can access services consistently and reliably, without interruptions or delays. Achieving and maintaining this quality is complex due to the dynamic, distributed, and resource-shared nature of cloud environments[10]. This background study explores the major concepts, existing challenges, and solutions in improving QoS in cloud systems.

QoS Metrics in Cloud Computing

To better understand how QoS can be improved, it is essential to identify the key performance indicators (KPIs) or metrics that define the quality of service in cloud computing environments:

Availability: The percentage of time a cloud service is accessible and operational. High availability ensures that services are always accessible to users.

Latency: The time taken to respond to a user request. Lower latency is crucial in performance-sensitive applications such as real-time data processing.

Reliability: The ability of a system to function without failure over a specific time period. High reliability is critical for long-running processes and transactions.

Throughput: The rate at which data is processed and transferred. Ensuring high throughput is necessary for data-intensive applications.

Scalability: The ability to handle increasing workloads by allocating more resources. A scalable system adapts efficiently to demand fluctuations.

Security and Privacy: Although security and privacy are often separate domains, their management affects QoS. Any breaches in security can degrade service quality and customer trust.

LITERATURE REVIEW

Wang, Jinjiang, *et al*[11] proposed strategy, termed LBVMP, seeks to establish a novel framework comprising a balanced flat surface of a physical machine (PM) regarding CPU, RAM, and bandwidth (BW), alongside another proportional flat surface representing the remaining resource capacity of the targeted PM divided by the requested resources (CPU, RAM, and BW) of a virtual machine (VM). Subsequently, LBVMP computes the distance between two plots to assess VM allocation solutions. Xu, Heyang, *et al*. [12] Examined the issue of fault tolerance-aware VM scheduling and articulated it as a multi-objective optimisation model incorporating various QoS constraints. The proposed model aims to minimise customers' overall spending while simultaneously maximising the successful execution rate of their enterprises. A greedy-based best fit decreasing (GBFD) algorithm is then designed to resolve the proposed optimisation model. The GBFD method employs a cost efficiency factor defined by the characteristics of CNs to select an appropriate CN for each VM request. Comprehensive experiments are performed to validate the practicality of the suggested models and algorithms using both real-world CDC cluster datasets and simulated data. Tamilarasu, P., and G. Singaravel [13] An Improved Coati Optimisation Algorithm-based Task Scheduling (ICOATS) is proposed to mitigate prolonged scheduling durations, excessive costs, and increased stress on Virtual Machines (VMs) in cloud computing environments. This suggested ICOATS constructs a model for work distribution and scheduling based on the variables of virtual machines, cost, and time. It also incorporated a multi-objective fitness function aimed for minimising makespan while simultaneously maximising resource utilisation efficiency. It established a potential strategy for each coati about the task scheduling process, which assists in identifying the appropriate assignment of incoming work to virtual machines (VMs). The proposal addresses premature convergence by integrating an exploitation method that enhances local search potential through a well-balanced trade-off between exploration and exploitation.

Rajak, Ranjit, *et al* [14] Task scheduling, characterised by dependencies between activities, is executed by resource allocation via Directed Acyclic Graph (DAG) scheduling. DAG is a crucial scheduling method because to its extensive applicability in various domains, including environmental technology, resource management, and energy optimisation. NP-completeness is a prominent issue, prompting the proposal of numerous models in the literature to address it. Nonetheless, the emergence of Quality of Service (QoS)-aware services in the CCE platform has become a significant and prevalent method for delivering computing resources, presenting a fresh essential challenge. The primary objective of this work is to formulate an innovative Directed Acyclic Graph (DAG) scheduling model to enhance the Quality of Service (QoS) parameters in the CCE platform, which can be validated using comprehensive simulation techniques.

Sharma, Minakshi, Rajneesh Kumar, and Anurag Jain [15] The proposed approach is an expansion of the previously suggested quality of service (QoS)-enabled join minimum loaded queue (JMLQ). The suggested methodology has been evaluated using the CloudSim simulator, and the findings indicate that it outperforms QoS-enabled JMLQ and its variants within the cloud context. Monika, and Om Prakash Sangwan [16] Utilised an innovative backpropagation-based Adaptive Dynamic Programming parameter tuning strategy, incorporating two fundamental prediction methods, to create a self-adaptive intelligent system that offers automatic parameter tuning capabilities for both techniques. To assess the suggested methodology, we conducted a simulation using a real QoS dataset, and the experimental findings indicate superior prediction accuracy relative to conventional methods.

Pakhrudin, Nor Syazwani Mohd, Murizah Kassim, and Azlina Idris [17] The aim of this research is to enhance the efficacy of the existing RR approach for action scheduling in the cloud by reducing the average waiting, turnaround, and response times. The CloudAnalyst tool was employed to refine the RR approach by adjusting parameter values to optimise for high accuracy and cheap cost. The results indicate that the total minimum and maximum response times achieved are 36.69 ms and 650.30 ms, respectively, for a duration of 300 minutes of RR. The expense for the virtual machines (VMs) ranges from $0.50 to $3.00. The duration of usage correlates positively with the expense of data transfer. This research is crucial for enhancing communication and the quality of interactions among groups.

Malla, Parvaz Ahmad, and Sophiya Sheikh [18] Cloud computing systems are recognised as significant consumers of energy resources globally. Moreover, power consumption has emerged as a critical factor since the majority of cloud computing systems rely on conventional nonrenewable energy sources. To render data centres environmentally sustainable, it is essential to implement optimal strategies to minimise energy usage and their detrimental impact on the environment. The primary purpose of this research is to examine several ways for constructing and sustaining an energy-efficient cloud. The paper will thoroughly examine several energy-efficient resource provisioning techniques and present a graphical comparison analysis of Quality of Service (QoS) metrics in cloud computing. Furthermore, the current study delineates the domains requiring enhancement to augment the energy efficiency of cloud computing systems.

Katkar, Alok, *et al* [19] This abstract provides an overview of current achievements in the automated assessment of quality of service in cloud computing systems. Cloud computing architectures provide substantial flexibility and scalability to clients. The total performance of a cloud platform is significantly affected by the quality of service. The quality of service is generally determined by outstanding criteria, including response time, availability, throughput, security, and others. Intelligent approaches have gained increased prominence for enhancing the quality of service monitoring and measuring. These tactics employ Artificial Intelligence (AI) and machine learning (ML) technologies to detect anomalies. Automated measures are implemented to guarantee the maintenance of superior service quality. Tabassum, Nazia, and C. R. K. Reddyy [20] VANET and Cloud Computing will significantly contribute to the advancement of efficient technology for autonomous driving, vehicle control, and intelligent systems in the near future. Cloud computing is a centralised paradigm that fails to adequately manage numerous Quality of Service (QoS) parameters, such as latency, throughput, and bandwidth optimisation. Fog Computing (FC) is established in VANETs to address the constraints of Cloud Computing (CC). The IoV-CC must tackle issues related to security and privacy. Consequently, the security protocols employed in conventional VANET and CC must be revised for IoV-CC, necessitating the development of a new secure algorithm to ensure secure communication between FOG and cloud nodes. Innovating QoS in VANET for IoV-CC faces considerable challenges related to data dissemination and security. This project aims to investigate the data distribution and security acceptability of the Internet of Vehicles (IoV) in relation to centralised and decentralised computer systems.

Arunkumar, J. R. [21] The computer resources of cloud service providers are reassigned dynamically based on demand, with their infrastructure, platform, software, and other resources shared among various corporate and private clients. The continuous rise of cloud computing subscribers utilising shared resources has heightened concerns regarding cloud security. This review article delineates present cloud security challenges and practices, while proposing several creative solutions aimed at enhancing cloud computing security in future domains. Agarwal, Rajesh, and Sanjay Dhingra [22] This research aims to identify the determinants of cloud service quality and evaluate the influence of service quality on customer satisfaction and loyalty. A study with 419 cloud experts/users was executed in India utilising a structured questionnaire based on the Likert scale. The participants were cloud professionals and users utilising the services of the five leading cloud service providers in India. Research hypotheses were evaluated by partial least squares structural equation modelling. The research indicated that agility, service assurance, dependability, scalability, security, service responsiveness, and usability all positively and significantly influence total cloud service quality. The study demonstrated a partly mediating impact of customer satisfaction between service quality and customer loyalty. Service quality exhibits a favourable and strong correlation with customer loyalty and customer satisfaction.

Pawar, Ankush Balaram, Shashikant U. Ghumbre, and Rashmi M. Jogdand [23] This work aims to create and construct a paradigm for authentication and data security in cloud computing. This technique comprises six distinct units: cloud server, data owner, cloud user, inspection authority, attribute authority, and central certified authority. The devised privacy preservation system has multiple stages: setup phase, key creation phase, authentication phase, and data exchange phase. The setup step is initially conducted by the owner, who provides the security attributes, while the key creation stage produces the system master key and the public parameter. Subsequently, the authentication process is conducted to ascertain the security measures of the information system. The data is

ultimately decrypted at the data sharing phase to facilitate data exchange and ensure privacy for confidential information. Furthermore, dynamic splicing is employed, alongside security mechanisms including hashing, Elliptic Curve Cryptography (ECC), Data Encryption Standard-3 (3DES), interpolation, polynomial kernel, and XOR, to safeguard sensitive data. Kirubakaran, S. Stewart, *et al* [24] Formulated a Privacy-Preserved Data Security Approach (PP-DSA) to provide data security and integrity for outsourced data in a Cloud Environment. This work ensures privacy preservation through the Efficient Authentication Technique (EAT), which use the Group Signature method in conjunction with a Third-Party Auditor (TPA). The auditor's responsibility is to safeguard data and ensure the integrity of shared information. Furthermore, the Cloud Service Provider (CSP) and Data User (DU) may also act as the perpetrators that must be addressed by the EAT. The primary aim of this effort is to improve cloud security and thus boost Quality of Service (QoS).

Sindjoung, Miguel Landry Foko, Mthulisi Velempini, and Clémentin Tayou Djamegni [25] Mobile Edge Computing (MEC) relocates computing and storage resources from cloud data centres to edge data centres, positioning them nearer to end-user devices to minimise end-to-end latency in request processing. Nonetheless, MEC is susceptible to security, data privacy, and authentication issues that impact the end-user Quality of Experience (QoE). It is essential to address these difficulties to prevent a subpar user experience resulting from inadequate security or data privacy. This research proposes a hybrid cryptographic system that integrates symmetric and asymmetric cryptographic methods to enhance data security, privacy, and user authentication in a MEC-based network. Guo, Zixuan, and Xuejun Yu [26] utilised the QoS (Quality of Service) of cloud services as the foundational data, derive the subjective weights of these services through AHP hierarchical analysis, ascertain the objective weights via the entropy weighting method, and calculate the recommendation degree for each service through weighting. Concurrently, they evaluate the trustworthiness of cloud services using the TOPSIS decision method, ultimately proposing the development of a cloud service trustworthiness metric model. This methodology mitigates the impact of user evaluation subjectivity on the trustworthiness assessment of cloud services by offering an objective and efficient trustworthiness metric. This study examines the processing of QoS data, the classification of metrics, and the application of the entropy weight approach to get the objective weights of the relevant metrics for cloud services.

Kaliyanandi, Maharajan, *et al* [27] developed a comprehensive strategy for load balancing in cloud computing that incorporates security measures. The Quantum-Based Security Framework has been developed, and the load is equilibrated by fuzzy logic. The primary security policies are effectively evaluated, and service is provided according to the user's specified requirements.The Security Framework devised a technique for cloud data storage by generating check bits in lieu of keys, enabling users to access their data upon verification of the check bits. Only the user may utilise the services and load balancing if the check bits produced by the user and the cloud service provider are same. Mirrored copies are created to mitigate the risk of data loss from failures or outages. This security technique enables us to achieve a high level of security. Liu, Xiaofei [28] The proposed methodology introduces a novel blended technique known as the Integrated Aquila Optimiser (IAO), which combines the traditional Aquila Optimiser (AO) with the Particle Swarm Optimisation (PSO) algorithm. The primary aim of this hybridisation is to address the deficiencies encountered by both AO and PSO algorithms. These algorithms are prone to becoming ensnared in local optima and exhibit restricted solution diversity. The suggested method introduces an innovative transition mechanism that enables appropriate changes between the search operators, assuring ongoing enhancements in the solutions. The transition method enables the algorithm to alternate between AO and PSO when either becomes stagnant or when solution variety diminishes. This adaptability improves the overall performance and efficacy of the hybrid method. The suggested IAO approach undergoes comprehensive testing via experiments done on the Cloudsim simulation platform.

Materwala, H., L. Ismail, and H. S. Hassanein [29] Introduced an innovative Artificial Intelligence QoS-SLA-aware adaptive genetic algorithm (QoS-SLA-AGA) to enhance application execution time for multi-request offloading in a heterogeneous edge-cloud computing environment, accounting for the effects of overlapping multi-request processing and variable vehicle speed. The suggested genetic algorithm incorporates an adjustable penalty function to accommodate the SLA constraints related to latency, processing time, deadlines, CPU, and memory needs.

Numerical investigations and analyses juxtapose our QoS-SLA-AGA with baseline genetic-based, meta-heuristic Particle Swarm Optimisation (PSO), random offloading, All Edge Computing (AEC), and All Cloud Computing (ACC) methodologies. Ali, Munwar, *et al* [30] Proposed a viable method to tackle these concerns with a novel service paradigm termed Confidentiality-based Classification-as-a-Service (C2aaS), which executes data processing by dynamically categorising data depending on its security level in anticipation of cloud storage. Our suggested service model demonstrates superior security for confidential data and effectively mitigates cloud system overloading compared to existing ways.

Research Gap

There are several inherent challenges in ensuring a high QoS in cloud environments due to their architecture and operational nature:

Dynamic Resource Allocation: Cloud environments operate on shared infrastructure, and resource contention can degrade QoS. As users' demands fluctuate, CSPs must allocate computing, storage, and network resources dynamically to ensure consistent performance. Poor resource allocation algorithms can lead to underutilization or resource overloading, both of which negatively affect QoS.

Geographically Distributed Data Centers: Cloud data centers are often distributed across different regions to minimize latency and ensure redundancy. However, geographic distribution introduces the complexity of managing network latency, data replication, and consistency across diverse locations. Ensuring uniform QoS in such scenarios is challenging, especially for global users.

Fault Tolerance and Failover: Faults in hardware, software, or network components can lead to service outages or reduced performance. High fault tolerance mechanisms are required to maintain QoS during failures. Failover strategies must be seamless and fast, ensuring no degradation in service during resource or network failures.

Multi-Tenancy and Resource Sharing: In a cloud environment, resources are shared among multiple tenants (users). While this leads to efficient resource utilization, it also introduces the risk of "noisy neighbors," where the actions of one tenant can negatively affect the performance of others. For example, a tenant running a heavy workload may consume more bandwidth or CPU resources, causing delays for other tenants.

Future Research Direction

Given these challenges, researchers and cloud service providers have proposed various strategies to improve QoS in cloud computing environments.

Dynamic Resource Allocation and Auto-Scaling: Dynamic resource allocation is one of the most effective strategies to ensure QoS. Cloud systems use **auto-scaling** mechanisms that dynamically allocate or deallocate resources based on real-time demand. By leveraging machine learning and predictive analytics, CSPs can forecast usage patterns and allocate resources more efficiently, preventing both resource underutilization and overloading.

Load Balancing Techniques: Load balancing is crucial for distributing workloads evenly across cloud servers. This helps in avoiding overloading specific servers, reducing response times, and maintaining high availability. Different algorithms are used to implement load balancing in cloud environments:

Round Robin: Distributes requests to all servers in a circular fashion.
Least Connections: Sends requests to the server with the fewest active connections, thereby preventing overload.
Dynamic Load Balancing: Uses real-time performance metrics to allocate resources based on current loads.

Service Level Agreements (SLAs):Service Level Agreements (SLAs) are contracts between CSPs and customers, outlining the expected QoS levels, including availability, response time, and performance guarantees. SLAs serve as a

framework for managing user expectations and hold CSPs accountable for performance and QoS failures. CSPs use SLAs to define specific thresholds, penalties, and compensations for non-compliance, incentivizing the maintenance of QoS.

Traffic Management and Network Optimization: Managing network traffic efficiently is essential to improving QoS in cloud systems. Network congestion or inefficient routing can lead to higher latencies and reduced throughput, negatively impacting performance. Advanced traffic management techniques, including Software-Defined Networking (SDN) and Network Function Virtualization (NFV), allow CSPs to dynamically manage network resources, optimize routing, and reduce latency.

REFERENCES

1. Tabassum, Nadia, *et al*. "Qos based cloud security evaluation using neuro fuzzy model." *Computers, Materials & Continua* 70.1 (2022): 1127-1140.
2. David, D. Stalin, *et al*. "Cloud Security Service for Identifying Unauthorized User Behaviour." *Computers, Materials & Continua* 70.2 (2022).
3. Parast, Fatemeh Khoda, *et al*. "Cloud computing security: A survey of service-based models." *Computers & Security* 114 (2022): 102580.
4. Kavitha, M. G., and D. Radha. "Quality, Security Issues, and Challenges in Multi-cloud Environment: A Comprehensive Review." *Operationalizing Multi-Cloud Environments: Technologies, Tools and Use Cases* (2022): 269-285.
5. Shanti, M. A., and K. Saravanan. "Knowledge data map—A framework for the field of data mining and knowledge discovery." *International Journal of Computer Engineering & Technology* 8.5 (2017): 67-77.
6. Shanti, M. A., and K. Saravanan. "An Effect of Data Mining Techniques in Public Healthcare-A Case Study." *International Journal of Civil Engineering and Technology* 9.9 (2018): 115-122.
7. Shanti, M. A. "A Study to Analyse the Quality of Work Life with Special Reference to Private Sector Bank Employees in Kumbakonam Town of Thanjavur District." *Our Heritage*, vol. 68, 2020.
8. Shanti, M. A., and K. Saravanan. "Wartortle—A Data Mining and Knowledge Discovery Suite for Data Analysis and Reporting." *International Journal of Applied Engineering Research*, vol. 13, no. 10, 2018, pp. 7835–7841.
9. Sahu, Parth, S. Raghavan, and K. Chandrasekaran. "Ensemble deep neural network based quality of service prediction for cloud service recommendation." *Neurocomputing* 465 (2021): 476-489.
10. Dheepak, T. "Enhancing the Cloud Security with ECC based Key Generation Technique." *Annals of the Romanian Society for Cell Biology* 25.2 (2021): 3874-3891.
11. Malathi, T., and T. Dheepak. "Enhanced Regression Method for Weather Forecasting." *The Scientific Temper*, vol. 15, special issue, 16 Oct. 2024, pp. 146–149. *The Scientific Temper*.
12. Suresh, T., T. Dheepak, and R. Kayalvizhi. "Optimizing QOS in Mobile Ad Hoc Networks Through Advanced Routing Protocols Under Wormhole Attack Scenarios." *International Journal on Recent and Innovation Trends in Computing and Communication*, vol. 11, no. 11, 2023, pp. 584–594.
13. Tamilarasu, P., and G. Singaravel. "Quality of service aware improved coati optimization algorithm for efficient task scheduling in cloud computing environment." *Journal of Engineering Research* (2023).
14. Rajak, Ranjit, *et al*. "A novel technique to optimize quality of service for directed acyclic graph (DAG) scheduling in cloud computing environment using heuristic approach." *The Journal of Supercomputing* 79.2 (2023): 1956-1979.
15. Dheepak, T. "Optimizing Routing Protocols in Mobile Adhoc Networks Using Firefly Optimization Algorithm." *Webology*, vol. 18, no. 5, 2021.
16. Dheepak, T. "Trust Based Cluster Selection for Intrusion Detection in Mobile Ad Hoc Networks." *Technology*, vol. 11, no. 10, 2020, pp. 421–430..

17. Dheepak, T. "Detection of Attacks in Wireless Networks Using Data Mining Techniques." *International Journal of Management (IJM)*, vol. 10, no. 5, Oct. 2019, pp. 280–288.
18. Ambika, G., and P. Srivaramangai. "Encrypted Query Data Processing in Internet Of Things (IoTs): CryptDB and Trusted DB." (2018).
19. Ambika, G. "Advanced Human Activity Recognition: Leveraging Adaptive Neural Networks and Diverse Machine Learning Algorithms on IoT Data." *Fuzzy Systems and Soft Computing*, vol. 19, no. 02(V), 2024, pp. 11–17.
20. G. Ambika, "IoTCryptDB: Encrypted query data processing in Internet of Things," *International Journal of Scientific Research in Computer Science Applications and Management Studies*, vol. 7, no. 4, p. 18, Jul. 2018.
21. Arunkumar, J. R. "Study Analysis of Cloud Security Chanllenges and Issues in Cloud Computing Technologies." *Journal of Science, Computing and Engineering Research* 6.8 (2023): 06-10.
22. Kasthuri, S., and A. Nisha Jebaseeli. "An artificial bee colony and pigeon inspired optimization hybrid feature selection algorithm for twitter sentiment analysis." *Journal of Computational and Theoretical Nanoscience* 17.12 (2020): 5378-5385.
23. Kasthuri, S., and A. Nisha Jebaseeli. "Study on social network analysis in data mining." *International Journal of Analytical and Experimental Modal Analysis (IJAEMA),(UGC CARE-A Journal), Impact Factor 6.3* 11.VIII (2019): 111-116.
24. Kirubakaran, S. Stewart, *et al*. "Towards Developing Privacy-Preserved Data Security Approach (PP-DSA) in Cloud Computing Environment." *Computer Systems Science & Engineering* 44.3 (2023).
25. Sindjoung, Miguel Landry Foko, Mthulisi Velempini, and ClémentinTayouDjamegni. "A data security and privacy scheme for user quality of experience in a Mobile Edge Computing-based network." *Array* 19 (2023): 100304.
26. Guo, Zixuan, and Xuejun Yu. "Cloud service quality assessment based on entropy weight method." *International Conference on Cryptography, Network Security, and Communication Technology (CNSCT 2023)*. Vol. 12641. SPIE, 2023.
27. Kaliyanandi, Maharajan, *et al*. "Design and development of novel security approach designed for cloud computing with load balancing." *AIP Conference Proceedings*. Vol. 2581. No. 1. AIP Publishing, 2023.
28. Liu, Xiaofei. "Hybrid Integrated Aquila Optimizer for Efficient Service Composition with Quality of Service Guarantees in Cloud Computing." *International Journal of Advanced Computer Science and Applications* 14.10 (2023).
29. Materwala, H., L. Ismail, and H. S. Hassanein. "QoS-SLA-aware adaptive genetic algorithm for multi-request offloading in integrated edge-cloud computing in Internet of vehicles. Vehicular Communications. 2023; 43: 100654." (2023).
30. Ali, Munwar, *et al*. "A Confidentiality-based data Classification-as-a-Service (C2aaS) for cloud security." *Alexandria Engineering Journal* 64 (2023): 749-760.