International Journal of Computer Science
Scholarly Peer Reviewed Research Journal - PRESS - OPEN ACCESS
ISSN: 2348-6600
http://www.ijcsjournal.com Volume 12, Issue 1, No 1, 2024. ISSN: 2348-6600
Reference ID: IJCS-490 PAGE NO: 3357-3369

AN ANALYSIS OF IMPROVED DYNAMIC LOAD BALANCING ALGORITHM IN CLOUD COMPUTING ENVIRONMENT

B.POONGODI
M.Sc Student,
Department of Computer Science,
Rajeswari College of Arts and Science for Women,
Bommayapalayam, Tamil Nadu, India.
Email ID: poongodipoongodi62313@gmail.com

Dr.A.SARANYA
Assistant Professor,
Department of Computer Application,
Rajeswari College of Arts and Science for Women,
Bommayapalayam, Tamil Nadu, India.
Email ID: drsaranyarcw@gmail.com

Abstract

One of the most crucial aspects of cloud computing in recent years has been regarded as load balancing. There are a lot more requests coming in quickly due to an increase in users worldwide. Global researchers have created a wide range of algorithms to process client requests at dispersed cloud servers. Accordingly, the cloud computing paradigm will automate server configuration for effective load balancing. Therefore, the load balancing method must be used to arrange the virtual machine selection process effectively. The availability of the virtual machine serves as the basis for the load balancing technique suggested in this work. Specifically, each virtual machine (VM) has its Availability Index (AI) value assessed over a predetermined duration; the AI value determines which work is assigned to that machine. Round Robin, Throttled, and Active Monitoring are three well-known load balancing algorithms that are compared with the suggested model in order to confirm its validity. With Cloud Analyst, the effectiveness of every algorithm was assessed. Comparing the suggested technique to alternative algorithms, simulation results demonstrate that it is more effective at load balancing between virtual machines.

Keywords: Cloud Computing, Modified Throttled, Virtual Machine, Throttled Algorithm, Round-Robin Algorithm, Active Monitoring.

I. Introduction

As a result of recent advances in Information and Communication Technology (ICT), computing is now regarded as a utility, similar to utilities like gas, electricity, water, and telephone service [1]. Customers can access those products whenever it suits their needs. Service providers receive payment from customers based on their consumption.

Similar to all other current utilities, computing utilities provide the fundamental computing services needed to meet the daily needs of the general public. Several computing paradigms, the most recent of which is called cloud computing, have been proposed to realize this objective. All that the cloud represents is a sizable collection of virtual resources that are readily available for use.

According to Dr. Rajkumar Buyya, a cloud is a kind of distributed, parallel system made up of several virtualized, networked computers that are dynamically provisioned and displayed as one or more merged computing resources in accordance with service-level agreements that exist between customers and service providers. [2] Cloud computing is a computer model that uses central, isolated servers and the internet to maintain apps and statistics. With the help of this technology, end users and enterprises can access apps on any computer with an internet connection without entering their personal information. Cloud computing centralizes memory, storage, distribution, and bandwidth, enabling far more efficient computation. A few instances of cloud computing are Google, Yahoo, Gmail, Hotmail, and so forth. The cloud service provider is in charge of managing both the email administration software and server, which are both hosted on the cloud. The final user enjoys the benefits and is able to utilize the software independently. In cloud computing, shared resources, software, and data are sent to computers and other devices, functioning more like a service than a product. Three categories can be used to group cloud computing services:

- ✓ SaaS (Software-as-a-Service)
- ✓ PaaS (Platform-as-a-Service)
- ✓ IaaS (Infrastructure-as-Service) respectively [3].

International Journal of Computer Science
Scholarly Peer Reviewed Research Journal - PRESS - OPEN ACCESS
ISSN: 2348-6600
http://www.ijcsjournal.com Volume 12, Issue 1, No 1, 2024. ISSN: 2348-6600
Reference ID: IJCS-490 PAGE NO: 3357-3369

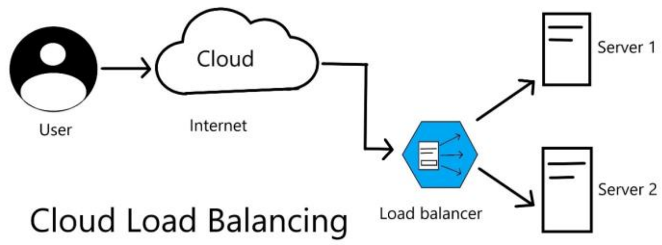

Fig 1: Cloud Load Balancing

Cloud Load Balancing: One of the most important methods in cloud computing for maximizing resource usage and making sure that no resource is overloaded with requests is load balancing. In order to improve speed, availability, and scalability, workloads are distributed among several computing resources, including servers, virtual machines, and containers.

User: Cloud storage stores data, including files, business data, movies, and photographs, on remote servers. Data is uploaded by users across the internet to servers, where it is stored in a virtual machine within a physical server.

Cloud: A distributed group of servers that host infrastructure and software is known as the cloud; it can be accessed online.

Internet: The Internet, or more precisely, whatever you can access remotely over the Internet, is the cloud, to put it simply. Anything that is kept on Internet servers rather than the hard drive of your computer is said to be in the cloud.

Load Balancer: In order to prevent any one server or system from being overloaded, underloaded, or idle, load balancing is used in cloud computing. In order to enhance overall cloud performance, load balancing optimizes a number of limited characteristics, including execution time, response time, and system reliability.

Server: Before forwarding requests to servers, load balancers do health checks on those servers. To prevent service interruptions and preserve high availability, load balancing automatically redirects the workload to a working server if one server is about to fail or is unavailable for maintenance or upgrades.

A. Cloud Software as a Service (SaaS)

Under this service model, the cloud customer consumes software services provided by the cloud provider that are hosted on cloud infrastructure, as opposed to locally executed applications. The software services that cloud consumers utilize are maintained and managed by the cloud provider. The cloud provider may impose fees based on the amount of software used and the duration of usage. The greatest approach to leverage cutting-edge technology is with SaaS. Examples of such service models are Customer Relationship Management (CRM) and Salesforge.com.

B. Cloud Platform as a Service (PaaS)

The cloud platform provides an environment for developers to construct and launch apps under this service framework. It offers a platform on which services and apps can operate. Although they are not obligated to handle the network, servers, operating system, or storage, customers retain control over the deployed applications. Examples of this type of approach are Right Scale, Microsoft Azure, and Google Application Engine [4].

C. Cloud Infrastructure as a Service (IaaS)

Cloud providers manage a huge variety of computing resources, including processing and storage capacity, under this service framework. The operating system, storage, installed apps, and sometimes limited access over specific networking components (such host firewalls) are all within the control of the cloud user. It is also referred to as hardware as a service (HaaS) at times. Here, the cost of the hardware can be significantly lowered. IaaS is provided by Amazon Web Services, Open Stack, Eucalyptus, GoGrid, and Flexiscale.

D. Cloud Resource Allocation

Resource allocation, resource scheduling, and resource supply are interchangeable phrases that have the following basic definitions: The process of assigning a service provider's resources to a client is known as resource supply. The process of distributing resources among equal groups of users or programs is known as resource allocation. Finally, resource scheduling is a plan of resource allocation in which computational activities are scheduled during intervals when resources are shared and made available. Put differently, it refers to the process of determining the start and finish times of computing processes based on their (1) predecessors, (2) previous relationships, (3) resource allocation, and (4) length. Moreover, resource search, selection, provisioning, application scheduling, and resource management are all included in cloud resource allocation.

Decisions about how much, what, when, and where to give the user access to the available resources are also part of the cloud resource allocation process (Fig 1 shows a block diagram of this process). Resource types and quantities are often specified by users, and in response, service providers assign the requested resource containers to their data centers. Applications must be executed efficiently if the type and quantity of resource

International Journal of Computer Science
Scholarly Peer Reviewed Research Journal - PRESS - OPEN ACCESS
ISSN: 2348-6600
http://www.ijcsjournal.com Volume 12, Issue 1, No 1, 2024. ISSN: 2348-6600
Reference ID: IJCS-490 PAGE NO: 3357-3369

containers meet certain requirements (such as a deadline for job completion) and are appropriate for the workload. Users of cloud computing environments can dynamically request or return resources thanks to the environment's elasticity; in this case, it's also important to think about how to implement such modifications.

Therefore, in order to provide effective cloud services and cloud-based applications, one must consider the traits and actions of actors in a cloud computing environment. When we say "efficient," we mean that the right resources are given to the right application at the right moment so that the application may make optimal use of them. Stated differently, effective resource allocation reduces the number of resources required for an application to maintain an acceptable level of service quality or maximizes production speed (or minimizes job completion time) of an application.

E. Significance of Resource Allocation

The process of assigning available resources via the internet to the necessary cloud applications in cloud computing is known as resource allocation (RA) [5]. Services will suffer if resource allocation is not carefully managed. Resource provisioning solves that problem by granting the service providers authority over the resources for each distinct module. In order to meet the demands of the cloud application, Resource Allocation Strategy (RAS) integrates cloud provider activities for the purpose of using and assigning scarce resources within the constraints of the cloud environment. It needs the kind and quantity of resources that each program needs to finish a task for the user. For an ideal RAS, the sequence and sequence of resource allocation are also inputs.

The following criteria should not be present in an ideal RAS:

- ✓ A process known as resource contention occurs when two apps attempt to use the same resource simultaneously.
- ✓ When there are few resources available, scarcity of resources occurs.
- ✓ An isolated state of resources leads to resource fragmentation. [6]
- ✓ Over-provisioning occurs when an application receives more resources than it requests.
- ✓ Under-provisioning of resources happens when there is not enough resources allocated to the application in comparison to the demand.

Oversupplying of resources may result from resource users' (cloud users') predictions of resource demands to finish a task before the estimated time. Under provisioning of

International Journal of Computer Science
Scholarly Peer Reviewed Research Journal - PRESS - OPEN ACCESS
ISSN: 2348-6600
http://www.ijcsjournal.com Volume 12, Issue 1, No 1, 2024. ISSN: 2348-6600
Reference ID: IJCS-490 PAGE NO: 3357-3369

resources could result from resource providers' resource allocation. Inputs for a RAS are required from cloud providers as well as users in order to address the previous differences. The application need and Service Level Agreement (SLA) are two of the main inputs to RAS from the view of the cloud user. The inputs needed from the other side to manage and distribute resources to host applications by RAS are the offerings, resource status, and available resources. Any ideal RAS's result must meet requirements for reaction time, latency, and throughput. Although cloud computing offers dependable resources, it also presents a significant challenge for dynamic resource allocation and management among apps.

It is impossible for a cloud provider to forecast the changing needs of users, applications, and users themselves. The task should be finished quickly and cheaply for cloud users. Therefore, we require an effective resource allocation system that works in cloud environments due to resource scarcity, heterogeneity, locality constraints, environmental requirements, and the dynamic nature of resource demand. The inputs needed from the other side to manage and distribute resources to host applications by RAS are the offerings, resource status, and available resources. Any ideal RAS's result must meet requirements for reaction time, latency, and throughput. Although cloud computing offers dependable resources, it also presents a significant challenge for dynamic resource allocation and management among apps.
It is impossible for a cloud provider to predict the changing needs of users, applications, and users themselves. The task should be finished quickly and easily for cloud users. We thus require an effective resource allocation system that works in cloud environments because of resource scarcity, heterogeneity, locality constraints, environmental requirements, and the dynamic nature of resource demand.

Both virtual and physical resources are included in cloud resources. Virtualization and supplying allow the physical resources to be shared among several compute requests. A collection of parameters describing the processor, memory, and disk needs are used to characterize the request for virtualized resources; this is shown in Fig 1. Through the mapping of virtualized resources to actual ones, supplying fulfills the request. Cloud applications are assigned hardware and software resources based on demand. Using virtual machines allows for scalable computing.

Large systems such as massive clusters, data centers, or grids have exponentially more complexity in determining the best resource allocation. Different solutions are recommended for resource distribution since

International Journal of Computer Science
Scholarly Peer Reviewed Research Journal - PRESS - OPEN ACCESS
ISSN: 2348-6600
http://www.ijcsjournal.com Volume 12, Issue 1, No 1, 2024. ISSN: 2348-6600
Reference ID: IJCS-490 PAGE NO: 3357-3369

the supply and demand of resources might be unexpected and dynamic. Several resource allocation techniques used in cloud systems are presented in this study.

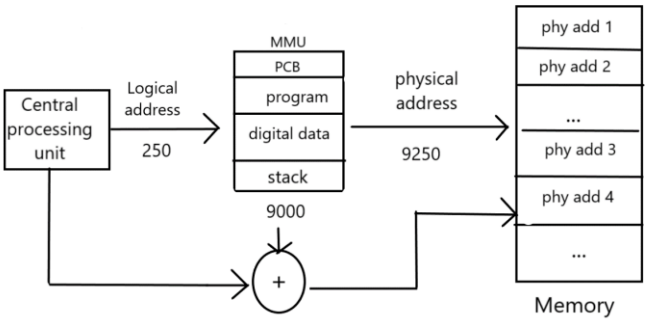

Fig 2: Mapping of Virtual to Physical Resources:

Physical Address:

A physical address, also known as a real address or a binary address, is a memory address that is used by the address bus circuitry to represent a binary number in order to allow the data bus to access a specific main memory storage cell or register of a memory-mapped I/O device.

Logical Address:

The address that a memory cell, storage component, or network host appears to be from the standpoint of an application program that is currently running is called a logical address. An address translator or mapping function's activity could cause a logical address to differ from a physical address.

CPU:

The physical host allots a time window for your vCPU to accept commands, carry out operations, and interact with other virtual components if it needs to run several vCPUs. Each virtual CPU is essentially a thread that shares a physical host's CPU cycle.

MMU:

The hardware part of a computer that manages all memory and cache functions related to the processor is called a memory management unit (MMU).

II. Applications

Applications connected to automotive touch range from simple transactions involving vehicle rank information to extremely complex large-scale visitor connections that include foundation integration. This service provides a summary of proposed request groups for vehicular networks as a starting point for examining requests. Even in cases when certain system functionalities aren't consistent for the highest volume of requests, this type of collection can never be finished entirely. Nonetheless, the evaluation highlights the elements, workings, and boundaries of the configuration. This gives an early grasp of the VANET touch residences and culminates in the next phase's more in-depth analysis of net aspects.

Safety Applications

Requests for safety can significantly lower the variety of traffic incidents. A driving force made up of a notification that arrives half

International Journal of Computer Science
Scholarly Peer Reviewed Research Journal - PRESS - OPEN ACCESS
ISSN: 2348-6600
http://www.ijcsjournal.com Volume 12, Issue 1, No 1, 2024. ISSN: 2348-6600
Reference ID: IJCS-490 PAGE NO: 3357-3369

a second before the collision is likely to prevent 60% of accidents, according to a few studies. There are three main conditions in which requesting protection is most likely beneficial.

1. Accidents

On major roads, cars travel at a faster speed. This gives the motive force a very short reaction time to the car in front of them. The approaching cars usually smash before they can come to a stop if an accident happens. Requests for protection are likely used to notify vehicles of an accident that occurred closer to the road, preventing a collision. In order to provide drivers with early warnings and prevent an accident from happening at the scene, a protection request may also be made.

2. Intersections

One of the most difficult tasks for drivers is navigating around and next to intersections since there are more traffic flows and a significant chance of collisions. In the United States, there were 9213 deaths that occurred at intersections, according to the U.S. Department of Transportation (DoT). These incidents made up more than 45% of all identified crashes and 21% of lives. The quantity of damages must drop if a safety request informed the motorist of an impending collision.. The motivating force may then take action to stop it.

3. Road Congestion

Additionally, it's likely that protection requests are made to give drivers the best routes to their destinations. This should increase the capacity of the roads and prevent interruptions in traffic caused by site visits by reducing street congestion and maintaining a level go with the flow of traffic. Because drivers would have to be less dissatisfied and more likely to follow visitors' instructions, it might also have a side effect of reducing traffic injuries.

4. User Applications

Consumers may be able to access information, entertainment, and ads through user requests while traveling. Below are two direct requests related to the consumer.

5. Internet Connectivity

It is now necessary for countless people to have regular access to the Internet, and since countless requests also require Internet connectivity, it is crucial to provide this capability to automobile passengers as well as extra VANET requests. Additionally, in this manner, the standard corporate structure might be present in cars without requiring any special reconstruction.

6. Peer-to-Peer Applications

Peer-to-peer queries are also an exciting idea for VANETs to reduce inactivity. Passengers in the cars can assign music, movies, and other media while conversing and playing all the extra fun games. On lengthy journeys, they can even stream music or videos from excellent servers.

International Journal of Computer Science
Scholarly Peer Reviewed Research Journal - PRESS - OPEN ACCESS
ISSN: 2348-6600
http://www.ijcsjournal.com Volume 12, Issue 1, No 1, 2024. ISSN: 2348-6600
Reference ID: IJCS-490 PAGE NO: 3357-3369

Finally, industrialized requests for VANETs should safeguard against those invisible setbacks that are inherent to VANETs. We see a touch resolution that was alerted in the following preparations.

III. Related Work

García, A.G., Blanquer, 2015 [7] A general framework for representing cloud services is proposed in this paper. This technology captures and manipulates arbitrary services using SLA pieces by utilizing the WS-Agreement specification. SLA segments are dynamically created in reaction to user requests. A prototype implementation of the methodology on a SLA-aware Cloud platform is made possible by a SLA composition algorithm. The flexibility, extensibility, and genericity of this technique enable the unification of Cloud service modeling. Ultimately, a use case offers a numerical assessment of the methodology's utility from the perspectives of cloud providers and users.

Pascual, J.A., Lorido-Botrán, T., Miguel-Alonso 2015 [8] Three multi-objective optimization methods with crossover and mutation operators particular to the problem have been put to the test. Simulation-based experiments show how a low-cost optimization leads to better resource assignments compared to traditional placement strategies, resulting in faster application execution and lower data center energy consumption. Both cloud customers and cloud providers benefit from this.

Singh, S., Chana 2015 [9] In particular, QoS (Quality of Service)-aware autonomic resource management and autonomic resource management in the cloud are covered in-depth and methodically in this research. There are several areas in which cloud computing's autonomic resource management currently falls. Several industrial and academic groups have conducted a methodical analysis of autonomous resource management in cloud computing and related approaches. Additionally, a taxonomy of cloud-based autonomous resource management has been provided. In addition to identifying key future research areas, this study will assist researchers in determining the salient features of autonomic resource management and in determining which technique is best for autonomic resource management in a given application.

Singh, S., Chana, I., Buyya 2016 [10] A new paradigm for effectively managing and delivering programs as services is cloud computing. New uses for cloud services are made possible by the convergence of cloud computing with technologies like big data analytics, wireless sensor networking, and the Internet of Things (IoT). This research presents an autonomous information system that runs on the cloud and uses big data and cloud computing to provide Agriculture-as-a-service (AaaS). The suggested system uses preset devices and Internet of Things sensors

International Journal of Computer Science
Scholarly Peer Reviewed Research Journal - PRESS - OPEN ACCESS
ISSN: 2348-6600
http://www.ijcsjournal.com Volume 12, Issue 1, No 1, 2024. ISSN: 2348-6600
Reference ID: IJCS-490 PAGE NO: 3357-3369

to collect data from a variety of users. Big data analytics is then used in the cloud to interpret the data, and consumers are automatically given the necessary information.. The suggested system's performance has been assessed in a cloud environment, and the results of the experiments indicate that it provides better service and has higher Quality of Service (QoS) in terms of QoS metrics.

Nguyen, Nguyen Cong 2017 [11] This research examines how adaptive algorithms and protocols for resource management in cloud networking might be developed using the economic and pricing models. Furthermore, we examine an array of incentive schemes based on pricing techniques for resource sharing in edge computing. Furthermore, we examine the application of pricing models in software-defined wireless networking that is cloud-based. In conclusion, we emphasize significant obstacles, unresolved concerns, and potential avenues for future study in the application of pricing and economic models to cloud networking.

Yousafzai, A., Gani, A., Noor 2017 [12] The virtues and weaknesses of the most recent state-of-the-art cloud resource allocation techniques are highlighted in this study through an exhaustive review. In addition, a theme taxonomy based on goals for resource allocation optimization is offered to categorize the body of current literature. The theme taxonomy is used to study the cloud resource allocation methods in order to identify their similarities and differences. Lastly, a number of options are put out for the creation of the best resource allocation plans.

Weerasiri, Denis, et al 2017 [13] Effective study, understanding, comparing, and choosing of cloud resource orchestration models, languages, platforms, and tools is made possible by this approach. This paper offers just that, offering a fresh and comprehensive analysis of the pertinent state-of-the-art in cloud resource orchestration.

Kumar, C. Ashok, et.al 2017 [14]A new technology called cloud computing allows users to access networks on demand by sharing cluster computing resources. Nonetheless, it can be challenging to keep multiple jobs running well in a cloud environment. As a result, it needs a load balancing strategy that distributes the workload among the Virtual Machines (VMs) without compromising system performance. The fractional dragonfly based load balancing algorithm (FDLA), which is based on the proposal of two selection probabilities, is a load balancing technique that is presented in this research. This work developed two selection probabilities, TSP and VSP, and proposed a load balancing technique, called FDLA, in the cloud computing system utilizing a novel load balancing algorithm, dubbed fractional dragonflies. The suggested load balancing methodology is predicated on the

International Journal of Computer Science
Scholarly Peer Reviewed Research Journal - PRESS - OPEN ACCESS
ISSN: 2348-6600
http://www.ijcsjournal.com Volume 12, Issue 1, No 1, 2024. ISSN: 2348-6600
Reference ID: IJCS-490 PAGE NO: 3357-3369

assessment of many factors, including machine loads and capacity. The load and number of tasks reallocated are used to evaluate the suggested FDLA approach in the cloud context. When the suggested FDLA is compared to three existing techniques—PSO, HBB-LB, and DA—it is able to achieve a minimal load of 0.2133 with a reallocation of 14 tasks, indicating that the technique is successful and performs significantly better.

Basu, Sayantani,et. al 2019 [15]In order to achieve virtualization in the cloud environment, resource usage and energy management must be properly managed. Load balancing is a crucial factor to take into account, as it distributes the workload to prevent any one node from becoming overworked. An improper load balancing configuration will result in memory and energy consumption losses. In a cloud, the resources ought to be scheduled so that users can access them whenever they want and perhaps with reduced energy waste. Every virtual machine has a node assigned to it. Every node's virtual machines are equivalent to a chromosome's genes. After completing crossover and mutation processes, the final work allocation was obtained by the application of optimization techniques. When it comes to load balancing and resource utilization, the suggested strategy has proven to be competent in comparison to previous approaches. It has been suggested to use an enhanced genetic algorithm with local search (GA-LS) to manage resource usage and load balancing. In this instance, the goal function was to lower the amount of memory and energy used by virtual machines that were assigned to cloud environment nodes.

Bhandari, Anmol et.al 2019 [17] In the recent past, load balancing has been regarded as one of the most crucial components of cloud computing. A significant number of requests are coming in at a quick pace due to an increase in users worldwide. Numerous algorithms have been developed by researchers worldwide to process client requests at dispersed cloud servers. In light of this, the cloud computing paradigm will automate server configuration to provide effective load balancing. From now on, the load balancing technique must be used to efficiently arrange the selection of virtual machines.. An algorithm for load balancing based on the virtual machine's availability is presented in this work. In particular, each virtual machine (VM) has its Availability Index (AI) value assessed over a predetermined duration; the AI value is then used to determine which task should be assigned to that machine. The suggested model is validated by comparing it to three well-known load balancing algorithms: Active Monitoring, Throttled, and Round Robin. Using Cloud Analyst, each algorithm's performance was assessed. According to simulation data, the suggested approach outperforms alternative algorithms in terms of load balancing between virtual machines.. The idea of task assignment in any

International Journal of Computer Science
Scholarly Peer Reviewed Research Journal - PRESS - OPEN ACCESS
ISSN: 2348-6600
http://www.ijcsjournal.com Volume 12, Issue 1, No 1, 2024. ISSN: 2348-6600
Reference ID: IJCS-490 PAGE NO: 3357-3369

circumstance—during compile time or at runtime—is taken into consideration by load balancing algorithms. Furthermore, it's critical to transfer the load to the proper virtual machine (VM) because the rise in user traffic has put a strain on the machine's computing capacity. Therefore, a modified throttled balancing technique is suggested in this study and is used with CloudSim's Cloud Analyst tool. Furthermore, it is verified through comparison with alternative load balancing strategies. The above comparison shows that the proposed load balancing algorithm is more effective and efficient as compared to other conventional load balancing algorithms. However, evaluation of appropriate load in real time still remains an open challenge for the future perspective.

IV. Conclusion and Future Scope

Both at the advanced and short levels, various methods for guaranteeing optimal resource allocation in cloud computing systems have been studied and examined. The text outlines the suggested solutions to overcome the obstacles to increasing the efficiency and speed of resource allocation. Despite some notable advancements in dynamic resource allocation speed increase, further progress remains to be made. Many problems are still unresolved, though. The continuous rise in processing power over the past 20 years has resulted in an unmanageable flow of data. This leads to the development of a clear difference between the volume of data being generated and the capacity of traditional tools to gather, process, and utilize this data most effectively. The financial benefits of cloud computing have made it quite popular in recent years. Carefully, cloud computing has offered the exploitations of demanding of data apps a number of benefits for its hosting. Modern cloud platforms have expanded the methods for more effectively allocating resources. However, a number of scheduling techniques have been created for flexible and effective resource distribution. Indeed, resource handling and accessibility—which directly impacts energy utilization—must be monitored in order to suitably ensure applications with QoS.

References

[1] Chana, I., Singh, S.,"Quality of service and service level agreements for cloud environments: issues and challenges", In: Cloud Computing-Challenges, Limitations and R&D Solutions, pp. 51–72. Springer International Publishing, 2014.

[2] Buyya, Rajkumar, et al.,"Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility," Future Generation computer systems 25.6, pp. 599-616, 2009.

[3] Weerasiri, Denis, et al.,"A Taxonomy and Survey of Cloud Resource Orchestration Techniques." ACM Computing Surveys (CSUR) 50.2 (2017): 26.

http://www.ijcsjournal.com Volume 12, Issue 1, No 1, 2024. ISSN: 2348-6600
Reference ID: IJCS-490 PAGE NO: 3357-3369

[4] García, A.G., Espert, I.B., García, V.H.,"SLA-driven dynamic cloud resource management", Futur. Gener. Comput. Syst. 31, pp. 1–11, 2014.

[5] Petcu, D.,"Consuming resources and services from multiple clouds", J. Grid Comput. 12(2), pp. 321–345, 2014.

[6] Singh, S., Chana, I.,"Formal Specification Language Based IaaS Cloud Workload Regression Analysis", arXiv preprint arXiv: 1402.3034. [Online] Available: http://arxiv.org/ftp/arxiv/papers/1402/1402.3034.pdf (2014)

[7] Szabo, C., Sheng, Q.Z., Kroeger, T., Zhang, Y., Jian, Y.: Science in the cloud: Allocation and execution of data- intensive scientific workflows. J. Grid Comput. 12(2), pp. 245–264, 2014.

[8] García, A.G., Blanquer, I.,"Cloud services representation using SLA composition. J. Grid Comput. 13(1), pp. 35–51, 2015.

[9] Pascual, J.A., Lorido-Botrán, T., Miguel-Alonso, J., Lozano, J.A.,"Towards a greener cloud infrastructure management using optimized placement policies", J. Grid Comput. 13(3), pp. 375–389, 2015.

[10] Singh, S., Chana, I.,"QoS-aware autonomic resource management in cloud computing: A systematic review", ACM Comput. Surv. 48(3), 39, 2015.

[11] Singh, S., Chana, I., Buyya, R.,"Building and Offering Aneka- based Agriculture as a Cloud and Big Data Service", Big Data: Principles and Paradigms, pp. 1–25. Elsevier 2016.

[12] Nguyen, Nguyen Cong, et al.,"Resource management in cloud networking using economic analysis and pricing models: A survey." IEEE Communications Surveys & Tutorials (2017).

[13] Yousafzai, A., Gani, A., Noor, R. M., Sookhak, M., Talebian, H., Shiraz, M., Khan, M. K.,"Cloud resource allocation schemes: Review, taxonomy, and opportunities", Knowledge and Information Systems, 50(2), pp. 347-381, 2017.

[14] Kumar, C. Ashok, R. Vimala, KR Aravind Britto, S. Sathya Devi,"FDLA: Fractional dragonfly based load balancing algorithm in cluster cloud model." Cluster Computing 22, No. 1, pp. 1401-1414, 2019.

[15] Basu, Sayantani, G. Kannayaram, Somula Ramasubbareddy, C. Venkatasubbaiah, "Improved Genetic Algorithm for Monitoring of Virtual Machines in Cloud Environment." In Smart Intelligent Computing and Applications, pp. 319-326. Springer, Singapore, 2019.

[16] Bhandari, Anmol, Kiranbir Kaur,"An Enhanced Post- migration Algorithm for Dynamic Load Balancing in Cloud Computing Environment." In Proceedings of International Ethical Hacking Conference 2018, pp. 59-73. Springer, Singapore, 2019.