Selected basic cloud load balancing algorithms and their comparative analysis

Srinivasa Rao G [1], Anuradha T [2]

Ph.D. Research scholar, Department of Computer Science, Dravidian University, AP, India [1] ,

Ph.D. Research Supervisor, Department of Computer Science, Dravidian University, AP, India [2] ,

Abstract*: Data centers are the backbone and they support service-oriented and deployment-oriented infrastructures in the cloud. There is always a quest for a good algorithm to properly schedule and maintain the resources on par with the requests and to balance the load. In this study, round-robin, throttled, equally spread current execution, artificial bee colony optimization, min-min, max-min, and htv algorithms are considered for the parameters, response time, throughput, virtual machine usage, energy usage, scalability, and complexity. Several other researcher's studies are taken into consideration and drawn a comprehensive conclusion as a round-robin algorithm has recorded below average (i)response time,(ii)scalability (iii)complexity (iv)throughput; but more (i)virtual machine usage,(ii)energy usage. The throttled algorithm has medium (i) response time, (ii) throughput, (iii) virtual machine usage; but it has attained more (i) energy usage (ii) scalability, below average complexity. The equally spread current execution algorithm has recorded medium (i)response time, (ii)throughput, (iii)virtual machine usage; But More energy usage, medium scalability, and below average complexity. The artificial bee colony optimization algorithm has good (i) response time, (ii) throughput; medium (i) virtual machine usage, (ii) energy usage; more (i) scalable,(ii)complexity. The min-min has good response time, more (i)throughput, (ii)scalable,(iii)complexity, medium (i)virtual machine usage, (ii)energy usage; The max-min algorithm has good response time, more throughput, medium (i)virtual machine usage, (ii)energy usage, more (i)scalability,(ii) complexity. The htv algorithm has medium (i)response time, (ii)throughput, (iii)virtual machine usage, more (i)energy usage,(ii)scalable, below average complexity.

Index Terms— Cloud computing, Utility computing, SaaS, PaaS, IaaS, SLA

I.INTRODUCTION

Cloud computing is an enticing field of computer science [1] for the information technology industry and also for the researchers. The father of cloud computing John Mc Karthy has a wish to have computing 'as- a- utility' and he aspired for the massive participation of the general public. His motto has been accomplished with cloud computing. Without massive storage infrastructure, cloud technology could not be successful.

One can access the cloud services using any of the devices and also their location is not at all a matter.This cloud technological advancement has improved the computational capability by reducing the restrictions such as device-oriented computing and location-based computing. This 'rented collection of node-based computing' [2] has been made to opt for many information technology industrial tycoons and start-ups as their best choice. This 'emerging engineering wonder' [3] has a great success due to the facility to adjust the configuration and to access larger distributed systems using the web, which no one has never dreamed of it. This 'business-oriented and service aimed' [4]cloud computing uses the 'highest theoretical background' [5]such as object-oriented computing. It has more 'practical and technological feasibility' [6].

Cloud is an 'interactive technology' [7] that provides the resources as per the needs of the user. In case of a service model (it should be called as a delivery model which is one of the working models of cloud computing) assures that the user would have the Software-as-a-service (SaaS) for his/her software requirements, Platform-as-a-service (PaaS) for applicational needs, and Infrastructure-as-a -service (IaaS) such for his/her infrastructures like virtual machines. These services are available interms of different blends of deployment models such as public, private, hybrid, and community models[8].The public cloud deployment model is intended for the general public usage, private deployment model for private usage and the hybrid deployment is a combination of both. The community model is entitled to use for a restricted community[9]. Working models may differ by any other means, but it is intended for the efficient utilization of computational resources along with reaching the end-user.

Cloud can be re-defined in terms of democratic terminology as, "for the end-user, by the service providers and to the end-user"[10]. The service provider will provide the service to the end-user based on Service Level Agreement(SLA). In order to provide the services effectively, the service provider has to manage his/her resources efficiently; else the user may contradict him/her for the violation of SLA. If such information goes viral into the cloud service user's community, that may cause the loss in their business [11]. There are a few issues observed with the implementation of the existing load balancing algorithms such as round robin in which the clients have to wait in the waiting queue un-till and un-less a suitable virtual machine is available, and the additional load on the scheduler to decide the size of quantum. In case of some other algorithms like throttled algorithm it works properly only if all virtual machines in data centre have the same hardware configuration[12]

II. NEED OF EFFICIENT USAGE OF RESOURCES

2.1. Cloud Resources utilization

On the practical grounds, in data centers, it is observed

four different types of challenges are arise quite commonly.

They are (i)at the infrastructure level, (ii)service providing level, (iii) performance and quality of service, (iv) growth of load on servers.

The infrastructural management of the data center is a big task. The factors like power, cooling, rack space and CPU Usage, Infrastructure components such as UPS devices, PDUs, chillers, HVACs, generators, branch circuits, etc. are a few to mention, under the best combination it is possible to provide good service[13].

Some of such issues like Electricity consumption depends on the size and the number of servers used. Data Center usually consumes 0.1 MW to 15 MW. Power Usage Effectiveness (PUE) is used to estimate the power usage in Data Centers [14].

Even though Moore's Law says about the increment of performance of transistors every approximately 730days, in case of power consumption there is no much improvement is evident. Apart from this Cloud Data Centers use 'Active' or 'Passive topology mode' to operates in disaster rescue mode. Green cloud computing, Energy-aware load balancing is also a new trend in Cloud Computing to reduce the energy consumption in Cloud Data Centers. Since the Data Centers emit much more heat, there is a need to cool to provide an environment-friendly data center.

CRAC units (Computer Room Air Conditioning Units) are used for the temperature maintenance, distribution of air in the Data Center, and the humidity level to control [15].

From the end-user (customer) point of view, problems are generally two types. They are, (i). Security to the data, (ii). Cost and the Efficient Services of the Cloud Services, (iii) Backup and recovery of data in case of loss of data. The end-user always feels his/her data should be safe and should not be accessed without any proper authentication process. Data leakage to the others is not accepted by any end-user, at the same time they need the cloud services at lower prices and efficient services. All the customers insist on the backup and recovery of data in case of loss of data. Cloud computing resources are a collection of computational resources and network resources. These resources are the real assets of cloud computing. Memory, CPU time, processing time, processing speed, disk storage, bandwidth, etc are a few to mention here. These computational resources usage can be measured partially by Big O notation. However, there is a need for proper planning for the distribution of cloud resources. Such plans and strategies are defined and mentioned in terms of an algorithm. The defined algorithm must have the fore vision to confirm the appropriate usage of computational usage. As the efficiency of the algorithm increases, usage will be optimized accordingly. Generally, an algorithm is engineered in such a way to optimized recurring processes or an optimized repetitive process of computational and network resources. Flexibility is needed in any algorithm to achieve improved results[16].

2.2. Resources allocation in cloud

Cloud computing is a collection of heterogeneous environments, thereby the nature of the user-requests also reflects the same. User requests need different computational requirements, and they vary from time to time. Peak hours and non-peak hours makes a good factor in predicting the resource requirements. During the peak, hours more resources need to be made available to handle the situation.

Resources need to be scheduled as per the needs. Scheduling algorithms existing numerous in literature and they have their advantages and disadvantages within their limits. Some algorithms good at complexity but fail to achieve scalability.round-robin algorithm has below average response time but more virtual machine usage. Throughput has a medium response time but it needs more energy. These all are in comparison with other algorithms studied in this study. It is impossible to provide a quality of service without the proper scheduling of computational resources [17].

Cloud services are provided to the end-users based on their service-oriented architecture(SOA). SOA is provided with the help of the internet. The success of cloud computing is possible with the help of virtualization technology. Virtualization helps to slice the data center to act like many numbers of servers. Even though it depends on the hardware configuration, there is a need for software to analyze and give a calculation of how many virtual machines are needed and how they can be divided to implement virtualization. Such types of software commercially available today to improve performance[18]for example Microsoft Azure.

The Quality of Services which is the most desired and agreed point of view of both service providers and the end-user. It depends on several points, and one of them is load balancing.

The load refers to the amount of data (traffic) being carried by the cloud computing network as a result of requests arrived from the end-users. The load balancing mechanism dependent on the amount of work allotted to the system for a specific period. Cloud computing is generally considered as on-demand service. Cloud load balancing algorithms play an important role in improving the performance of cloud service providing and for the overall resource utilization [19].

The objective of the service provider is to provide maximum resources output; it is possible to achieve by implementing a better load balancing algorithm. To distribute the workload of multiple network links, to achieve maximum throughput, to minimize the response time, and to avoid overloading, proper scheduling is to be made appropriately[20]. Proper Scheduling of tasks would reduce the problem of starvation. There are many load balancing algorithms of different types and categories, but all of them had their limitations. Although it is not possible to make the best load balancing algorithm, improvement in this regarding is always an open choice[21].

In computation, a 'job' is a collection of a set of tasks. A 'task' is a piece of code that would be used inside a thread execution. 'Execution time' and processing time would play an important role in the aspect of computation [22]. Execution time is a time that is actively utilized by a particular job or an activity using the micro processor's computing resources. 'Response time' is a time in which a job or an activity becomes activated. Response time would be longer than the execution time in an operating system [23].

Generally, scheduling is a term used for a set of rules that will govern the order by the execution of a particular process. Here the process is a word used for an executable programming part. A process execution will undergo in the computer system such as a CPU burst followed by an IO burst again a CPU burst followed by an IO burst and so on. 'Burst' is an action with as fast as possible the highest speed and continued action. Scheduling is an important property of an operating system[24]. The scheduling process can be looked at as 'Service request scheduling' and 'Resource response scheduling'. Generally, service request scheduling occurs due to (i) When the user submits his or her request to the service provider (ii) Service provider executes the request (iii) Processing the request in the service request architecture (iv). Dynamic virtual machine generation and dispatch at the provider site [25]. A computer system needs scheduling before use; CPU is one of the most critical parts of the computer. Multiprogramming is one of the basic and important scheduling techniques. Generally, CPU scheduling is done in such a way as to keep it busy as much as possible [26].

Scheduling occurs due to any one of the four conditions such as the given below (i) When a process has gone from switching state to waiting for state or (ii) When a process is terminated or (iii) When a process switches from running state to the ready state or When a process switches from waiting for the state to the ready state. In the cases (i) and (ii) there is no choice for scheduling and this is non -pre-emptive scheduling. Here a new process is to be selected for the execution. It is also called as cooperative scheduling [27]. Under this scheduling, if one Central Processing Unit is allocated to a process, the process keeps the Central Processing Unit busy until it releases the Central Processing Unit or by terminating or by switching it to the waiting state. In the cases (ii) and (iii) there is a choice for scheduling. This is pre-emptive scheduling [28]. It is related to share data when a kernel on behalf of the processor is working and busy, 'processor waiting' is required. Rarely it could be interrupted, but it is not desirable.

Scheduling performance can be evaluated using a synthetic workload below the parameters (i).Distribution of several tasks, (ii). Distribution of real-time job services on demand, (iii).Average inter-arrival time of tasks, (iv).Period of real-time jobs.

Scheduling policies can be stated as below (i).FCFS (First Come First Serve), (ii).Round Robin (Execution of tasks in a cyclic manner), (iii). Shortest Job First (SJF) etc.

2.3. Cloud Resources and virtualization

Virtualization is the key enabler to cloud computing. It manages the abstraction of physical resources like virtual machines. In cloud computing server virtualization provides flexibility. Server virtualization is the best application for hardware virtualization. It changes the traditional concept of a single host. In the traditional system, most of the virtual machines left underutilized. Through virtualization it is possible to run multiple virtual machines on a single physical system and sharing the resources of one system to multiple is possible through server virtualization. Server virtualization can be used for different virtual machines with different operating systems on the same physical system. It is virtualization which is both one to many and many to one concept[29].

Virtualization provides a way to configure a collection of computing resources for different types of applications. This configuration is possible to make changes at the hardware level. Virtualization provides better services as it allows virtual machine dynamic migration to implement QFR (Quick Failure Recovery) without any interruption in the service to the user. Virtualization in cloud computing provides computing resources, which can be monitored and easily maintained. Resource monitoring can be automated. If at all any common computing resources are existing, then they can be reused [30].

Virtualization provides the applications to migrate from one server to another. This migration is possible dynamically, which means there is no need to make the server down, and the workload can be easily managed. Virtualization provides a way to run multiple applications at the same time and on the same server. In this way the effective way of computing resource utilization is possible.

Computing tasks are distributed in cloud computing. These tasks are mapped to a pool of resources. In a single computer system, it can create many numbers of virtual machines. If any host needs, it could be easily transferred its working load to another host with below average interruption. This type of transfer is called a virtual machine migration. In earlier days this process used to require shutting down the present working virtual machine and, allocating sufficient computing resources to the new virtual machine and then start working on the new virtual machine. With the help of live virtual machine migration, it does require. Here the downtime is very below average in live migration. Downtime is a time in which the virtual machine will not provide its service. At the time of performing the live virtual machine migration, there are two parameters in this regard. They are,

(i). Downtime: is a smaller amount of time in which a computer system would not be successful to make available or carry out its major functions.

(ii). Migration time: is a time taken to transfer, a particular virtual machine from a source to the destination host without any interruption in the service. Virtualization

partitions the computing resources into multiple units and then run them simultaneously and makes it share among the other virtual machines. The virtual machine migration is a process of transferring a file or a task from one virtual machine to another virtual machine.

Generally, virtual machine migration can be done in any of the two methods. They are,

(i). Non live virtual machine migration: It is also called a cold virtual machine migration. In this type of migration virtual machine is needed to down and the services will not be available for a considerable amount of time.

(ii). Live Virtual machine migration: It is also called a hot virtual machine migration. In this type of migration virtual machine is not needed to down and the services are available without any interruption. Live migration is used for load balancing. Using this type of migration it can be optimized the virtual machine development in the data centers. Live virtual machine migration can be achieved in any one of two types. They are,

(a).A control mechanism is switched to the destination.
(b).Data is transferred to the destination: this data transfer can be achieved in any one of the following:

(i). Pre copy: This type of copy allows transferring the memory to the destination first and the execution is transferred later.

(ii). Post copy: This type of copy allows us to transfer the execution to the destination first and then the memory later. Generally, the cloud user needs the tasks to be completed with below average makespan (completion time of the given job), with below average computational cost, below average turnaround time, and with below average response time. At the same time, the cloud provider needs below average resource utilization, more throughput, energy efficiency, and a balanced load.

The incoming 'user request traffic load' needs to be balanced. The balanced load will make the resources to be ready as par with the requirements and makes it to utilize effectively and the cost of computation to be reduced and increases the Quality of Service to the end-user.

Static Load Balancing is accomplished by providing prior information approximately about the system resources [31]. The overall performance of the node is determined at the time of execution. Nodes calculate their allotted works and submit the result to the remote node. Then relying on the overall performance workload is sent at the being without considering the current status of the system. Static load balancing strategies are non-pre-emptive i.e. As soon as the load is allotted to the node it cannot be transferred to another node. This approach requires below average communication, as a result, reduces the execution time. The major drawback of this method is that it does not take the present status of the gadget even at the time of making allocation choices. This has the principal impact on the overall performance of the system due to load fluctuation in a distributed system. For example, there are a few static load balancing algorithms are available such as round-robin, central manager, threshold algorithm, and randomized algorithm [32].

In Dynamic Load Balancing the Present status of the system is considered this algorithm and then load distribution is done accordingly. This set of rules is generally composed of three techniques: transfer strategy, location strategy, and information strategy. For instance, Central Queue Algorithm, stores new requests in a FIFO (first in first out) queue. Each new request is inserted in the queue. If a high prioritize request is there then it would give a chance to that prioritized request for processing. If there isn't any requested activity inside the queue then the request is buffered till a new activity is to be had [33].

III. EXISTING BASIC LOAD BALANCING ALGORITHMS AND AN ANALYSIS BASED ON EXPERIMENT (METHODOLOGY FOLLOWED)

Today we have many numbers of load balancing algorithms. Here are a few basic algorithms used widely and they have their limitations.

This present paper has dealt with load balancing algorithms. The main motto of any algorithm is to provide reduce resource utilization and to provide effective service to the end-user.

1. Round Robin algorithm (Basic version) is the easiest algorithm used for the logic division of time and nodes. This algorithm has a simple process with the automatic mechanism of selecting the next process in the queue. In this algorithm, the overall time is partitioned into the number of segments, and each node in the system is allocated with a particular time segment or time. Therefore it works on the logic of time division. Every node will be assigned with some time along with work. The main responsibility of a cloud service provider is accepting the request from the client, and allotting time segment for each request. But this Algorithm is static and does not maintain the present status of each node at the time of allocation [34].

2. Throttled algorithm (Basic version) : (i).It is a dynamic algorithm or static scheduling algorithm, (ii). It is a virtual machine-based algorithm, (iii). Throttled virtual machine load balancer keeps a list of virtual machines and their status whether they are busy or idling when the user cloudlet request comes assigns to a proper virtual machine and lets the work to be done [35].

3. Equally, spread current execution (ESCE) load balancing algorithm (basic version) distributes the load randomly by evaluating the size and transfers the load to that virtual machine which is lightly loaded or handles that task easy and takes below average time and gives increase throughput. In this spread spectrum technique load balancer spreads the jobs with multiple virtual machines. Generally,

the Basic version of the ESCE algorithm is Dynamic, having average throughput, average response time, average scalability, high priority, low fault tolerance, average cost overhead, high power consumption, low complexity, average fairness, and average performance.

4. Artificial bee colony optimization Algorithm basic version was inspired by the above-mentioned behavior of the bee swarm. Artificial Bee Colony Optimization Hybrid Algorithm is a novel metaheuristic approach that was developed in 2005. It is motivated by the intelligent foraging behavior of honey and is the simulation of the minimalistic foraging model of the honey bee in the search process for solving real-time parameter, non-convex, and non-smooth optimization problem. In this algorithm, a colony consists of three types of artificial bees including employed, onlooker, and scout bees. (i). It is a dynamic algorithm, (ii). Best Suited for cloud Environment, (iii). Best suited for Parallel Processing, (iv). There will be three types of Bees (a). Scout Bees: Arbitrarily searches for Honey, (b). Onlooker Bees: Determines the food source by calculating the fitness, (c).Employed Bees: Gather the information about the food resource and exchange the information gathered with onlooker bees.

Min-Min Algorithm: The algorithm will begin with a collection of tasks that are not assigned to any tasks to do. At first, below average time required tasks are selected and then the least minimum value is considered. As per the least time the task is scheduled.

Max-Min Algorithm: This algorithm follows similar to the Min-Min Algorithm, but instead of choosing the min value, the maximum value is chosen [36].

HTV Balancing Algorithm with Increased Priority: In this algorithm, tasks are performed effectively. The major points it covers are a load of the server, server statistics, and also the time constraints. This algorithm would perform the tasks efficiently in the data centers.

To check any algorithm, it should be tested in any simulation environment. Cloud analyst is a simulation environment. It is a better choice to test by providing different types of parameters for many number of times.

Let us denote the performance of various load balancing algorithms with the numbers as ‘below average performance’ - 1, ‘medium performance’ - 2, ‘good performance’ - 3, ‘high performance’ - 4.

Fig. 1: Cloud Analyst Simulation tool [37]

SL. No	Reg. Name	R0	R1	R2	R3	R4	R5
1	**R0**	**2000**	1000	1000	1000	1000	**1000**
2	**R1**	**1000**	800	1000	1000	1000	**1000**
3	**R2**	**1000**	1000	2500	1000	1000	**1000**
4	**R3**	**1000**	1000	1000	1000	1000	**1000**
5	**R4**	**1000**	1000	1000	1000	500	**1000**
6	**R5**	**1000**	1000	1000	1000	1000	**2000**

Table. 1. Available bandwidth between the regions measured in Mbps.

SL. No	Region Name	R0	R1	R2	R3	R4	R5
1	**R0**	20	10	10	10	10	10
2	**R1**	10	80	10	10	10	10
3	**R2**	10	10	25	10	10	10
4	**R3**	10	10	10	10	10	10
5	**R4**	10	10	10	10	50	10
6	**R5**	10	10	10	10	0	20

Table. 2. Delay matrix and transmission between the regions in nano sec.

SL. No	Region Name	R0	R1	R2	R3	R4	R5
D.C -1	**R0**	20	10	10	10	10	10
D.C -2	**R1**	10	80	10	10	10	10
D.C -3	**R2**	10	10	25	10	10	10
D.C -4	**R3**	10	10	10	10	10	10
D.C -5	**R4**	10	10	10	10	50	10
D.C -6	**R5**	10	10	10	10	10	20

Table.3. Data center configuration details

ID No	Availbl. Mem.	Storg. Capacit	Availble B/W	Number of Process.	Proces. Speed	Virtual machine policy

1	04800 B	100,000,000 GB	1,000,000 Mbps	4	10,000 Mbps	Time-shared Policy

Table.4. Physical hardware details

Name of the Scenario	No.of User bases	Regions	Req. user/hr	Req.size	Peak Hr start at GMT	Peak Hr end at GMT	Avg. No of Peak Users	Avg. No of OFF Peak Users
Scenario -1	1	0	60	20	3	9	1000	100
Scenario -2	3	2,4,5	60	20,30	3	9	1000	100
Scenario -3	6	1,2,4	60	1,100, 20,20	3	9	1000	100
Scenario -4	4	1,3,4	60	1,100, 20,30	3	9	1000	100
Scenario -5	5	1,2,5	60	1,10, 20,20	3	9	1000	100
Scenario - 6	3	0-5	60	1,20, 40,50	3	9	1000	100

Table.5. User base properties

Algorithm	Response time	Throughput	Virtual machine usage	Energy usage	Scalability	Complexity
Round -robin	1	1	4	4	1	1
Throttled	2	2	2	4	4	1
Equally spread current execution	2	2	2	4	2	1
Artificial bee Colonyoptimization	3	3	2	2	4	1
Min -min	3	4	2	2	4	4
Max-min	3	4	2	2	4	4
Htv	2	2	2	4	4	1

Table.6. Output table with Numbers assigned for the quality of parameters indicated as 1-Below Average,2-Medium, 3-Good,4-High.

The performance of 'response time' is depicted in the below Figure.2. Max-Min, Min-Min, and ABC algorithms have shown improved performance.

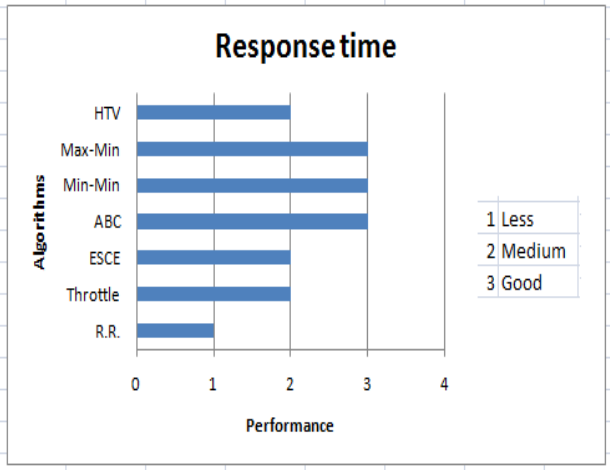

Fig.2. Different Load balancing algorithms and Response time.

The performance of 'Throughput' is depicted in the below Figure.3. Max-Min and Min-Min algorithms have shown improved performance in comparison with other algorithms. The performance of Throttled,ESCE,HTV have medium response time, and round robin has the lowest response time in comparision with other considered algorithms

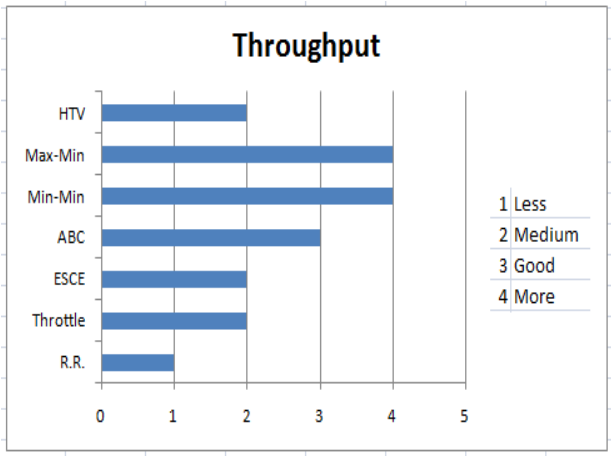

Fig.3.Different Load balancing algorithms and Throughput.

The performance of 'VM usage' is depicted in the below Figure.4. Round robin algorithm has shown improved performance, but ESCE, ABC,HTV has medium throughput,Min-Min,Min-Max has more throughput.

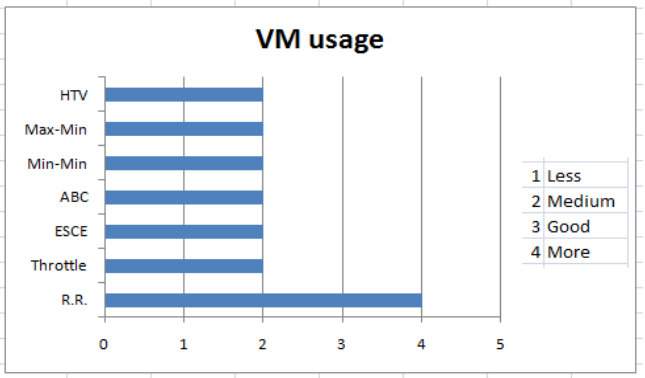

Fig.4. Different Load balancing algorithms and VM Usage.

The performance of 'Energy usage' is depicted in the below Figure.5. Max-Min and Min-Min algorithms have shown below average performance, but other algorithms have shown improved performance.

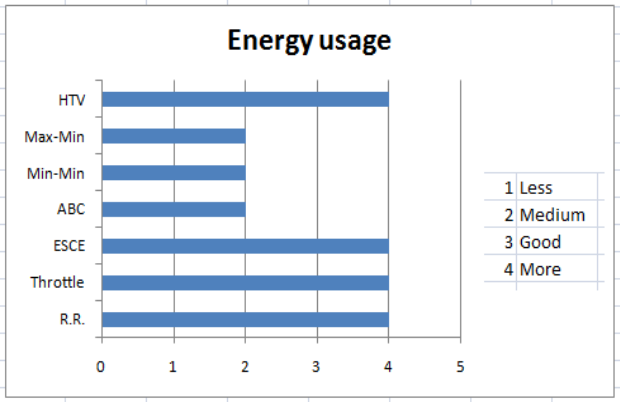

Fig.5. Different Load balancing algorithms and Energy Usage.

The performance of 'Scalable' is depicted in the below Figure.6. ESCE and Round-robin algorithms have shown below average performance, but other algorithms have shown improved performance.

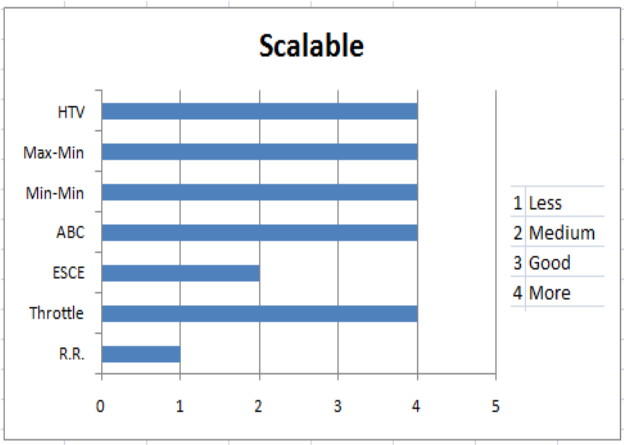

Fig.6. Different Load balancing algorithms and Scalable.

The performance of 'Complexity' is depicted in the below Figure.7. Max-Min, Min-Min, and ABC algorithms have shown improved performance, but other algorithms have shown below average performance.

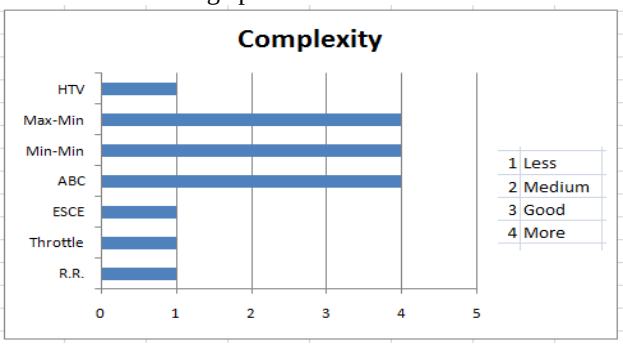

Fig.7. Different Load balancing algorithms and complexity.

IV. RESULTS AND DISCUSSION

From the above graphs, it is evident that Artificial Bee Colony Optimization and Min-Min have more response time and the Round robin algorithm has below average response time. Round robin usage of the virtual machine is more in comparison with that of the other algorithms. Min-Min and Max-Min algorithms have more throughput and Round-robin has fewer throughputs. Round Robin, Throttled, Equally Spread Current Execution Algorithm, and HTV algorithms have more Energy usage, and Artificial Bee Colony Optimization, Min-Min, Max-Min algorithms have below average Energy usage. In scalability point of view Throttled, Artificial Bee Colony Optimization, Min-Min, Max-Min algorithms are more scalable and Round robin, Equally Spread Current Execution algorithm are below average scalable. Artificial Bee Colony Optimization, Min-Min, Max-Min algorithms have more complexed and Round robin, Equally Spread Current Execution and HTV algorithms have below average complexity.

V. SOLUTION TO THE EXISTING PROBLEM

Till now many researchers have contributed for the improvement of load balancing in cloud computing environment. The process of load balancing need to consider many background needs such as whether it is needed to perform with in a given fixed time or it is needed to be virtual machine friendly, or it needs with more through put, or it needs more scalability, or it needs more energy. One single algorithms would not be able to provide proper scheduling mechanism in all the needed directions. Therefore 'hybrid algorithms' would be best suited to provide improved services to needs of the present conditions.

Some efforts are done by some researchers as given below.

1. Gill Sukhjindersingh, ThaparVivek proposed HACOBEE set of rules that is a premier solution for load balancing [37] most of the nodes. It showed better response instances.

2. Foram F Kherani, Prof. JigneshVania added a new type of approach called "Active monitoring load balancing Algorithms (AMLBA)". This approach continues track of all

virtual machines status and requests are allotted to the digital device popularity and requests are allotted to the digital gadget on FCFS approach. Once the request is finished its system load balancer [38] de allocates the digital machine.

3. Mohd Z. Freaj and Azzam Sleit proposed [39] a new mechanism to solve the trouble of useful resource provisioning. This useful resource provisioning is made the usage of a way known as "Hybrid Approach for resource provisioning in cloud computing (HARP)". This approach is created primarily based thresholding, statistical median, and virtual system multiplexing ideas. This results in computing assets and strength utilization.

VI. CONCLUSION

Cloud computing is gaining more popularity among the industry and start-ups. 'It is a subscription-based renting mechanism' of computational resources. The computational resources being allocated as per the needs of the end-user. In this study the round-robin algorithm has recorded below average (i)response time,(ii)scalability (iii)complexity (iv)throughput; but more (i)virtual machine usage,(ii)energy usage. The throttled algorithm has medium (i) response time, (ii) throughput, (iii) virtual machine usage; but it has attained more (i) energy usage (ii) scalability, below average complexity. The equally spread current execution algorithm has recorded medium (i)response time, (ii)throughput, (iii)virtual machine usage; But More energy usage, medium scalability, and below average complexity. The artificial bee colony optimization algorithm has good (i) response time, (ii) throughput; medium (i) virtual machine usage, (ii) energy usage; more (i) scalable,(ii)complexity. The min-min has good response time, more (i)throughput, (ii)scalable,(iii)complexity, medium (i)virtual machine usage, (ii)energy usage; The max-min algorithm has good response time, more throughput, medium (i)virtual machine usage, (ii)energy usage, more (i)scalability,(ii) complexity. The htv algorithm has medium (i)response time, (ii)throughput, (iii)virtual machine usage, more (i)energy usage,(ii)scalable, below average complexity.

RECOMMENDITION

For smaller data centers and below average traffic load static load balancing algorithms are best suited but among them, some other parameters also need to consider such as more power consumption. This is one step towards establishing greed clouds. Meanwhile, if the algorithm takes more computational resources then it must be addressed with some improvement or by considering another algorithm.

FUTURE SCOPE

There are many algorithms that are there for the improvement of load balancing in cloud computing, but all of them have their own drawbacks. Different circumstances need different load balancing approaches. One single algorithm would not be able to provide proper scheduling mechanism in all the needed directions. Therefore 'hybrid algorithms' would be best suited to provide improved services to the needs of the present conditions. This present paper analyzed a few and this can be carry forwarded to the next level with some other algorithms and parameters with vigorous efforts to analyze.

ACKNOWLEDGEMENT

I sincerely thank my Research Supervisor and Vice-Chancellor (i/c) of Dravidian University, Prof.T.Anuradha for her elderly contribution to this paper and for her guiding through her teaching, research guidance which made me explore more into the qualitative content about the cloud computing environment. As since in this era, the world has become very gregarious for handling impeccable knowledge to savvy and scrupulous educators, my research supervisor has made me aim to stretch to gravitate innovative ideas in the field of cloud computing and this is a small effort to tees off and endorse to bring out improvements within my limit. She has thrived and successfully transformed my draft ideas into this format of this paper. It is all her inspiration and shown committed path which has given me to bring out this paper.

FUNDING STATEMENT

This Study is not funded by any organization

CONFLICT OF INTEREST

I donot have any conflict of Interest

REFERENCES

[1] Gundu, S.R., Panem, C.A. & Thimmapuram, A. Intelligence Using Automata-Based Nature's Digital Philosophy. SN COMPUT. SCI. 1, 189 (2020).

[2] Zhou, S. Wang, Q. Sun, H. Zou, and F. Yang, "A Dynamic Virtual Resource Renting Method for Maximizing the Profit of Cloud Service Provider under SLA Constraint," presented at the 2013 IEEE 6th International Conference on Cloud Computing (CLOUD), Jun. 2013, doi: 10.1109/cloud.2013.13

[3] Gundu, S.R., Panem, C.A. & Thimmapuram, A. Real-Time Cloud-Based Load Balance Algorithms and an Analysis. SN COMPUT. SCI. 1, 187 (2020).

[4] M. Karakus, L. Zengxiang, C. Wentong, and T. N. B. Duong, "OMTiR: Open Market for Trading Idle Cloud Resources," presented at the 2014 IEEE 6th International Conference on Cloud Computing Technology and Science (CloudCom), Dec. 2014, doi: 10.1109/cloudcom.2014.102.

[5] Xu Xiao-tao, Yang Chen, Dai Guang-hua, and Ma Hua-long, "Information service quality evaluation study of cloud computing environment based on big data," presented at the 2017 IEEE 2nd International Conference on Cloud Computing and Big Data Analysis (ICCCBDA), Apr. 2017, doi: 10.1109/icccbda.2017.7951917.

[6] Gundu, S.R., Panem, C.A. & Timmapuram, A. Robotic Technology-Based Cloud Computing for Improved Services. SN COMPUT. SCI. 1, 190 (2020).

[7] F. Pomerleau, F. Colas, and R. Siegwart, "A Review of Point Cloud Registration Algorithms for Mobile Robotics," FNT in Robotics, vol. 4, no. 1, pp. 1–104, 2015, doi: 10.1561/2300000035.

[8] P. R. da Cunha, P. Melo, and C. F. da Silva, "Avoiding Lock-In: Timely Reconfiguration of a Virtual Cloud

Platform on Top of Multiple PaaS and IaaS Providers," presented at the 2013 IEEE 6th International Conference on Cloud Computing (CLOUD), Jun. 2013, doi: 10.1109/cloud.2013.36.

[9] Gundu, S.R., Panem, C.A. & Thimmapuram, A. "The Dynamic Computational Model and the New Era of Cloud Computation Using Microsoft Azure". SN COMPUT. SCI. 1, 189 (2020). doi.org/10.1007/s42979-020-00276-y

[10] N. Nikolov and O. Nakov, "Research of Communication between STM32L475 and Private Cloud realized by using Amazon FreeRTOS and MQTT," presented at the 2019 27th National Conference with International Participation (TELECOM), Oct. 2019, doi: 10.1109/telecom48729.2019.8994890.

[11] [11].N. Bennani, C. G. Guegan, M. A. Musicante, and G. V. Solar, "SLA-Guided Data Integration on Cloud Environments," presented at the 2014 IEEE 7th International Conference on Cloud Computing (CLOUD), Jun. 2014, doi: 10.1109/cloud.2014.130.

[12] T. Deepa and D. Cheelu, "A comparative study of static and dynamic load balancing algorithms in cloud computing," presented at the 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS), Aug. 2017, doi: 10.1109/icecds.2017.8390086.

[13] R. Takahashi et al., "Torus Data Center Network With Smart Flow Control Enabled by Hybrid Optoelectronic Routers [Invited]," J. Opt. Commun. Netw., vol. 7, no. 12, p. B141, Oct. 2015, doi: 10.1364/jocn.7.00b141.

[14] [14].M. Xu, C. Liu, and S. Subramaniam, "PODCA: A Passive Optical Data Center Network Architecture," J. Opt. Commun. Netw., vol. 10, no. 4, p. 409, Mar. 2018, doi: 10.1364/jocn.10.000409.

[15] M. O. Onibonoje, P. N. Bokoro, N. I. Nwulu, and S. L. Gbadamosi, "An IoT-Based Approach to Real-Time Conditioning and Control in a Server Room," presented at the 2019 International Artificial Intelligence and Data Processing Symposium (IDAP), Sep. 2019, doi: 10.1109/idap.2019.8875880.

[16] J. Hayes, "Data centres 2020," Engineering & Technology, vol. 4, no. 19, pp. 54–54, Nov. 2009, doi: 10.1049/et.2009.1912.

[17] P. Maenhaut, H. Moens, B. Volckaert, V. Ongenae and F. De Turck, "Resource Allocation in the Cloud: From Simulation to Experimental Validation," 2017 IEEE 10th International Conference on Cloud Computing (CLOUD), Honolulu, CA, pp. 701-704, doi: 10.1109/CLOUD.2017.96, 2017.

[18] Salapura, "Cloud computing: Virtualization and resiliency for data center computing," 2012 IEEE 30th International Conference on Computer Design (ICCD), Montreal, QC, pp. 1-2, doi: 10.1109/ICCD.2012.6378606, 2012.

[19] Gundu, S.R., Panem, C.A. & Thimmapuram, A. "The Dynamic Computational Model and the New Era of Cloud Computation Using Microsoft Azure". SN COMPUT. SCI. 1, 189 (2020). doi.org/10.1007/s42979-020-00276-y

[20] Laghari, A. A., He, H., Khan, A., Kumar, N., & Kharel, R. (2018). Quality of experience framework for cloud computing (QoC). IEEE Access, 6, 64876-64890.

[21] Rashid Nazir, Zeeshan Ahmed, Zeeshan Ahmad, Noor Nabi Shaikh, Asif Ali Laghari, Kumlesh Kumar, Year: 2020, Cloud Computing Applications: A Review, CS EAI DOI: 10.4108/eai.22-5-2020.164667

[22] [22].Kumar, V., Laghari, A. A., Karim, S., Shakir, M., & Brohi, A. A. (2019). Comparison of fog computing & cloud computing. Int. J. Math. Sci. Comput, 1, 31-41.

[23] Halepoto, I. A., Memon, M. S., & Parveen, S. (2017). Analysis of quality of experience frameworks for cloud computing. IJCSNS, 17(12), 228.

[24] Shafiq, M., & Khan, A. (2018). Assessment of quality of experience (QoE) of image compression in social cloud computing. Multiagent and Grid Systems, 14(2), 125-143.

[25] Gundu, S.R., Panem, C.A. & Thimmapuram, A. "Hybrid IT and Multi Cloud an Emerging Trend and Improved Performance in Cloud Computing". SN COMPUT. SCI. 1, 189 (2020). doi.org/10.1007/s42979-020-00277-x

[26] Laghari, A. A., He, H., Shafiq, M., & Khan, A. (2017, May). Impact of storage of mobile on quality of experience (QoE) at user level accessing cloud. In 2017 IEEE 9th international conference on communication software and networks (ICCSN) (pp. 1402-1409). IEEE.

[27] Memon, K. A., Laghari, R. A., Halepoto, I. A., & Khan, A. (2019). Quality of experience (QoE) in cloud gaming models: A review. Multiagent and Grid Systems, 15(3), 289-304.

[28] Laghari, Asif Ali, Hui He, Muhammad Shafiq, and Asiya Khan. "Assessing effect of Cloud distance on end user's Quality of Experience (QoE)." In 2016 2nd IEEE international conference on computer and communications (ICCC), pp. 500-505. IEEE, 2016.

[29] Laghari, A., He, H., Laghari, R., Khan, A., & Yadav, R. (2019). Cache Performance Optimization of QoC Framework. EAI Endorsed Transactions on Scalable Information Systems, 6(20).

[30] Wagan, A., & Umrani, A. (2020). Effect of Packet Loss and Reorder on Quality of Audio Streaming. EAI Endorsed Transactions on Scalable Information Systems, 7(24).

[31] T. Deepa and D. Cheelu, "A comparative study of static and dynamic load balancing algorithms in cloud computing," 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS), Chennai, pp. 3375-3378, doi: 10.1109/ICECDS.2017.8390086, 2017.

[32] Roopali Gupta, Meenu Dave, "Load Balancing Issues and Techniques In Cloud Computing", International Journal of Computer Sciences and Engineering, Vol.8, Issue.1, pp.137-140, 2020.

[33] Duraksha Ali, Manoj Kumar Gupta, "Deadline Sensitive Lease Scheduling Using Hungarian Genetic Algorithm in Cloud Computing Environment", International Journal of Computer Sciences and Engineering, Vol.7, Issue.12, pp.7-15, 2019.

[34] Sanjay G. Patel, S.D. Panchal, "QoS Parameters for Cloud Using Swarm Optimization Dynamic Load

Balancing Algorithm," International Journal of Scientific Research in Network Security and Communication, Vol.7, Issue.3, pp.33-39, 2019.

[35] U.Kaur, M. Mahajan, D. Singh, "A Comparative Analysis of Trust Models in Cloud Computing," International Journal of Scientific Research in Network Security and Communication, Vol.6, Issue.2, pp.19-23, 2018.

[36] Harish C. Sharma, Meenakshi Bisht, "Best Fit Resource Allocation in Cloud Computing", International Journal of Computer Sciences and Engineering, Vol.7, Issue.3, pp.871-875, 2019.

[37] Gill Sukhjinder Singh, ThaparVivek, "Implementation of A Hybrid Load balancing Algorithm for Cloud computing", 2nd International Conference on Science, Technology and Management, University of Delhi,27th Sep 2015 ,Pages: 173-182.

[38] Ankush P. Deshmukh, Prof. KumarswamyPamu, "Applying Load Balancing: A Dynamic Approach", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 2, Issue 6, June 2012, Pages: 201-206.

[39] Azzam Sleit1, Moh'd Z. Freaj,"Hybrid Approach for Resource Provisioning in CloudComputing", Scitech Research Organization-Volume 6, Issue 1| July 07, 2016 , Pages: 546|-561.