Chalanati Durga Venkata Peddintla Kumari* et al.
(IJITR) INTERNATIONAL JOURNAL OF INNOVATIVE TECHNOLOGY AND RESEARCH
Volume No.7, Issue No.4, June-July 2019, 9190-9196.

Accumulating Source Exploitation of Virtual Machinefor Load Balancing in Cloud Computing

CHALANATI DURGA VENKATA PEDDINTLA KUMARI
M.Tech Scholar (CSE) and Department of Computer Science Engineering, Kakinada Institute of Engineering and Technology for Women, Korangi, AP, India.

S V KRISHNA REDDY
Assistant Professor, Department of Computer Science and Engineering, Kakinada Institute of Engineering and Technology for Women, Korangi, AP, India

Abstract: **Load balancing in cloud computing has assumed a pioneer job in improving the effectiveness. From 10 years prior there has been a speedy progression in the use of web and its applications. Appropriated computing is generally called web based computing where we rent the enrolling resources over the web. It is a remuneration for every usage show where you pay for the proportion of organizations rented. It gives different central focuses over the customary computing. With cloud computing expanding such a colossal vitality now days, the working environment culture is despite changing a similar number of people now particularly wants to work from home rather than going every day to office. There are three essential organizations gave by cloud that are SAAS, IAAS and PAAS. Load balancing is an incredibly main problem faced now days in cloud condition with the goal that the benefits are capably utilized. There are many load balancing algorithms available that are used to adjust the load of the client requests. In this paper we will propose a methodology which is a mix of Honeybee Foraging Algorithm, Active clustering algorithm and Ant Colony Optimization.**

Keywords: **Cloud Computing; Load Balancing; Programing Management; Virtualization Machine;**

I. Introduction

Cloud computing is an innovation dependent on appropriated computing which causes clients to use computing assets on the web. Cloud gives benefits in pay as you use style. As cloud based administrations are used by clients over the globe, it is basic to have instruments to improve asset use through balancing load. This is accomplished utilizing a 'Load Balancer'. Cloud servers experience unexpected barges in solicitations. So as to process such asks for, it is important to have dynamic asset designation and furthermore load balancing arrangements. At the point when more assets are apportioned, it brings about wastage of assets. At the point when less assets are dispensed, it causes server quality issues or failure to serve customer demands. In this manner a hearty load balancing component is necessitated that aides in optimization of asset usage. The current algorithm proposed investigated load balancing issue in cloud computing. They utilized a system that utilizes cloud allotments and the status data, for example, IDLE, NORMAL AND OVERLOADED so as to settle on load balancing choices at server side. Solicitations are prepared by the server dependent on the parity of load on the servers. It means that demands that originate from client are sent to various parcels dependent on the load level of the servers associated with the segments. There is no arrangement for cloud parcel rules. Cloud servers with related load balancer can help in advancing the load balancing process in cloud computing. The invigorate period is an issue. Further examination is required so as to discover best revive period to refresh the load status. No advancement technique for load balancing. We executed a load balancing model where each parcel knows about load balancing model. In view of the status data the load balancing choices are made. The status data gives required load degree which is critical to settle on all around educated choices. Among various allotments and servers, load balancer projects will help in realizing coordination for settling on load balancing choices progressively at run time. The servers which are less loaded are given more demands while the servers which are overloaded are not given further demands. These choices are verifiably made by load balancers as they do have insight and measurements used to keep up load in ideal design. Network computing is forerunner of cloud computing [1], [2]. Nature of Service (QoS) is one of the necessities that are related with load balancing in cloud computing [3]. With load balancing set up, the servers in the cloud can perform better along these lines improving accessibility of cloud and adaptability. Versatility alludes to the capacity to serve expanded number of customers without corrupting execution. Accessibility alludes to the measure that shows the accessibility of server in a given year. These measurements likewise help in surveying generally speaking execution of the cloud. The open cloud considered in this task is fit for improving execution as far as idleness, CPU usage and speedup. Innovations assume fundamental job in the cloud computing worldview in future [4]. Asset provisioning can be improved in cloud utilizing certain benchmark rehearses as investigated in [5]. Cloud assets are helpful in applications related High Performance Computing (HPC). In the current answer for cloud load balancing proposed

in, there is no arrangement for cloud parcel rules. Cloud parcel principles can help in upgrading the load balancing process in cloud computing. The invigorate period is an issue. Further examination is required so as to discover best revive period to refresh the load status. There is no development system for load balancing. In this manner, another algorithm for cloud parceling is required so as to enhance the exhibition of cloud.The new algorithm should deal with cloud segment rules and furthermore perfect revive period which will deal with optimization issue in load balancing and asset usage. Invigorate period can serve the cloud as impetus for deciding perfect condition of the segments and settle on all around educated choices relating to load balancing. This paper deals with optimization of load balancing and client occupation planning.

II. Related work

Yongfei Zhu, Di Zhao [03], In the zone of cloud computing load balancing, the Particle Swarm Optimization (PSO) algorithm is neoteric and now lauded very, yet as of late a more neoteric algorithm which conveys the classifier into load balancing is displayed. In addition, an algorithm called red-dark tree which is going for improving the proficiency of asset dispatching is likewise commended. In any case, the 3 algorithms all have various disadvantages which can't be overlooked. For instance, the dispatch effectiveness of PSO algorithm isn't fulfilling; despite the fact that classifier and red-dark tree algorithm improve the productivity of dispatching undertakings, the exhibition in load balancing isn't that great, therefore the improved PSO algorithm is introduced. A few scientists are intended to get the advantages of new algorithm. Most importantly, the time multifaceted nature and execution for every algorithm in principle are registered; and after that real information which are created in investigations are given to exhibit the presentation. What's more, from the investigation result, it very well may be discovered that for the speed of algorithm itself PSO is the most reduced, and the improved PSO take care of this issue in some degree; improved PSO algorithm has the best execution in undertaking tackling and PSO is the subsequent one, the red-dark and Naive Bayes algorithm are much slower; PSO and improved PSO algorithm perform well in load balancing, while the other two algorithms don't progress nicely. GeethaMegharaj, Dr. Mohan K.G [04], Load Balancing is the one of the most significant parts in circulated situations. As Cloud Computing is outstanding amongst other stage that gives stockpiling of information in negligible expense and open forever over the web, load balancing for the cloud computing has transformed into an extremely intriguing and important investigation region. Load balancing supports to get a high client fulfillment and use of asset proportion by guaranteeing a capable and sensible allotment of each computing asset. There are various troubles in the load balancing strategies, for example, security, adaptation to non-critical failure and so forth in cloud computing conditions. Numerous scientists have been proposed a few methods to upgrade the load balancing. This paper depicts a review on load balancing plans in cloud conditions. We investigate the various sorts of algorithms that is proposed by numerous specialists to take care of the issue of load balancing in cloud computing. RenGao and Juebo Wu [05], How to appropriate and organize undertakings in cloud computing is a difficult issue, so as to get ideal asset usage and evade overload. In this paper, we present a novel methodology on load balancing by means of ant colony optimization (ACO), for balancing the workload in a cloud computing stage powerfully. Two systems, forward-in reverse ant instrument and max-min rules, are acquainted with rapidly discover the competitor hubs for load balancing. We detail pheromone introduction and pheromone update as indicated by physical assets under the haze computing condition, including pheromone dissipation, motivator, and discipline rules, and so forth. Joined with errand execution expectation, we characterize the moving likelihood of ants in two different ways, that is, regardless of whether the forward ant meets the retrogressive ant, or not, in the neighbor hub, with the point of quickening looking through procedures. Reenactments represent that the proposed technique cannot just give dynamic load balancing to cloud computing with less looking through time, however can likewise get high system execution under medium and intensely loaded settings.

III. Load Balancing

Load balancing is the process of improving the performance of the system by shifting of workload among the processors. Workload of a machine means the total processing time it requires to execute all the tasks assigned to the machine. Balancing the load of virtual machines uniformly means that anyone of the available machine is not idle or partially loaded while others are heavily loaded. Load balancing is one of the important factors to heighten the working performance of the cloud service provider. The benefits of distributing the workload includes increased resource utilization ratio which further leads to enhancing the overall performance thereby achieving maximum client satisfaction.

In cloud computing, if users are increasing load will also be increased, the increase in the number of users will lead to poor performance in terms of resource usage, if the cloud provider is not configured with any good mechanism for load balancing and also the capacity of cloud servers

would not be utilized properly. This will confiscate or seize the performance of heavy loaded node. If some good load balancing technique is implemented, it will equally divide the load (here term equally defines low load on heavy loaded node and more load on node with less load now) and thereby we can maximize resource utilization. One of the crucial issue of cloud computing is to divide the workload dynamically.

Load balancing is used to distributing a larger processing load to smaller processing nodes for enhancing the overall performance of system. In cloud computing environment load balancing is required to distribute the dynamic local workload evenly between all the nodes.

A. *Load balancing classification:*

Fig.3 represents different load balancing algorithms. This is mainly divided into two categories: static load balancing algorithm and dynamic load balancing algorithm:

1) **Static approach**: - This approach is mainly defined in the design or implementation of the system. Static load balancing algorithms divide the traffic equivalently between all servers.

2) **Dynamic approach**:- This approach considered only the current state of the system during load balancing decisions. Dynamic approach is more suitable for widely distributed systems such as cloud computing. Dynamic load balancing approaches have two types .They are distributed approach and nondistributed(centralized) approach. It is defined as following:

a) **Centralized approach**: - In centralized approach, only a single node is responsible for managing and distribution within the whole system. Other all nodes are not responsible for this.

b) **Distributed approach**: - In distributed approach, each node independently builds its own load vector. Vector is collecting the load information of other nodes. All decisions are made locally using local load vectors. Distributed approach is more suitable for widely distributed systems such as cloud computing.

B. *Metrics for Load Balancing*:

1. Throughput: - It is used to calculate the all tasks whose execution has been completed. The performance of any system is improved if throughput is high.

2. Fault Tolerance: -It means recovery from failure. The load balancing should be a good fault tolerant technique. 3. Migration time: -It is the time to migrate the jobs or resources from one node to other nodes. It should be minimized in order to enhance the performance of the system.

4. Response Time: - It is the amount of time that is taken by a particular load balancing algorithm to response a task in a system. This parameter should be minimized for better performance of a system.

5. Scalability: - It is the ability of an algorithm to perform Load balancing for any finite number of nodes of a system. This metric should be improved for a good system.

C. Policies of load balancing algorithm

There are many policies are used in load balancing algorithms:

- Information policy: It defined that what information is required and how this information is collected. This is also defined that when this information is collected
- Triggering policy: This policy defined that time period when the load balancing operation is starting to manage the load.
- Resource type policy: This policy defined the all types of resources which are available during the load balancing.
- Location policy: This uses all the results of the resource type policy. It is used to find a partner for a server or receiver.
- Selection policy: This policy is used to find out the task which transfers from overloaded node to free node.

D. Major goals of load balancing algorithms

1. Cost effectiveness: Load balancing help in provide better system performance at lower cost.

2. Scalability and flexibility: The system for which load balancing algorithms are implemented may be change in size after some time. So the algorithm must handle these types' situations. So algorithm must be flexible and scalable.

3. Priority: Prioritization of the resources or jobs needs to be done. So higher priority jobs get better chance to execute.

Goals of Load Balancing

Goals of load balancing include:

- Substantial improvement in performance
- Stability maintenance of the system
- Increase flexibility of the system so as to adapt to the modifications.
- Build a fault tolerant system by creating backups.

Classification of Load Balancing Algorithm

Based on process orientation they are classified as:

a) Sender Initiated: In this sender initiates the process; the client sends request until a receiver is assigned to him to receive his workload

b) Receiver Initiated: The receiver initiates the process; the receiver sends a request to acknowledge a sender who is ready to share the workload

c) Symmetric: It is a combination of both sender and receiver initiated type of load balancing algorithm.

Based on the current state of the system they are classified as:

1. Static Load Balancing

In the static load balancing algorithm the decision of shifting the load does not depend on the current state of the system. It requires knowledge about the applications and resources of the system. The performance of the virtual machines is determined at the time of job arrival. The master processor assigns the workload to other slave processors according to their performance. The assigned work is thus performed by the slave processors and the result is returned to the master processor.

Static load balancing algorithms are not pre-emptive and therefore each machine has at least one task assigned for itself. Its aims in minimizing the execution time of the task and limit communication overhead and delays. This algorithm has a drawback that the task is assigned to the processors or machines only after it is created and that task cannot be shifted during its execution to any other machine for balancing the load. The four different types of Static load balancing techniques are Round Robin algorithm, Central Manager Algorithm, Threshold algorithm and randomized algorithm.

2. Dynamic Load Balancing

In this type of load balancing algorithms the current state of the system is used to make any decision for load balancing, thus the shifting of the load is depend on the current state of the system. It allows for processes to move from an over utilized machine to an under-utilized machine dynamically for faster execution.

This means that it allows for process pre-emption which is not supported in Static load balancing approach. An important advantage of this approach is that its decision for balancing the load is based on the current state of the system which helps in improving the overall performance of the system by migrating the load dynamically.

Traditional Computing V/S Cloud Computing Environment

There are many similarities as well as differences between traditional scheduling algorithms and the scheduling of VM resources in cloud computing environment.

First of all the major difference between cloud computing environment and traditional computing environment is the target of scheduling. In traditional computing environment, it mainly schedules process or task so the granularity and the transferred data is small; whereas in cloud computing environment, the scheduling target is VM resources so the granularity is large and the transferred data is large as well.

Secondly, in cloud computing environment, compared with the deployment time of VMs, the time of scheduling algorithm can almost be neglected.

Need of Load Balancing

We can balance the load of a machine by dynamically shifting the workload local to the machine to remote nodes or machines which are less utilized. This maximizes the user satisfaction, minimizing response time, increasing resource utilization, reducing the number of job rejections and raising the performance ratio of the system.

Load balancing is also needed for achieving Green computing in clouds [5]. The factors responsible for it are:

1. Limited Energy Consumption: Load balancing can reduce the amount of energy consumption by avoiding over hearting of nodes or virtual machines due to excessive workload.

2. Reducing Carbon Emission: Energy consumption and carbon emission are the two sides of the same coin. Both are directly proportional to each other. Load balancing helps in reducing energy consumption which will automatically reduce carbon emission and thus achieve Green Computing.

IV. Load Balancing Workflow

Load balancing [7]-[9] is a computer networking technique to disseminate workload throughout several computers or network links, a computer cluster, disk drives, central processing units or other resources to obtain optimal resource utilization, minimize response time, maximize throughput and ignore overhead. Utilizing several workflows with load balancing, rather than a single workflow may reduce efficiency of the system by scalability. The load balancing service is normally offered by devoted hardware or software, i.e. a system server or a multilayer switch. In this research paper we have introduced the some of the load balancing mechanisms which targets the effective usage of the cloud computing resources [10]. First one is the normal load balancing algorithm in which if more no. of subscribers or requests (for example, n = 50) to the specific virtual machine greater than the load balancing server redirect the incoming request to the other virtual machine. The other kind of load balancing algorithm is the load balance depending on the existing disk space or memory space on the virtual machine. Most cloud subscribers utilize the cloud

resources either to save or fetch data for some calculation. Load Balancing Operational Workflow The assignment procedure is a mathematical programming mechanism which is utilized to detect the optimal solution for requested problems.

Fig. 1 illustrates the operational workflow for introduced mechanism.

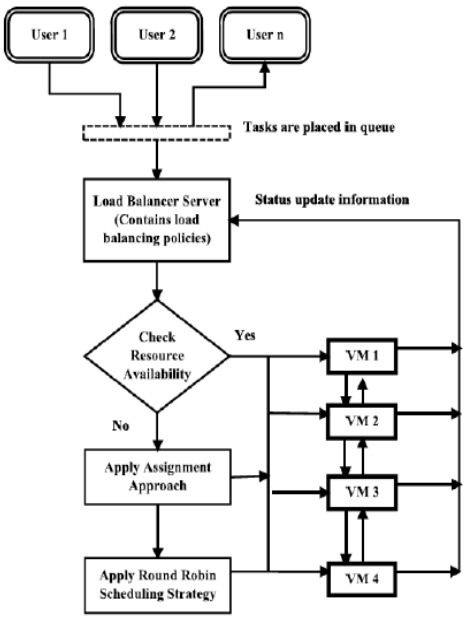

Figure 1: Load Balancing Operational Workflow

The above diagram describes the complete workflow or operation of this paper. For this research paper we have made a web application in which „n" no. of subscribers is permitted to access the virtual machines effectively from the client side. First the different subscribers utilize the application and send the task to the server. The subscribers requests are positioned in the task queue and it is forwarded to the load balancer. The load balancer examines for the existed virtual resources which are linked to it and it also has the status information of each virtual machine which is linked to it. The status information shows the status of each virtual machine i.e. whether it is in busy state or in presented state. Depending on the virtual machines status the task can be assigned to the virtual resources. If no. of tasks is existed to no. of existed virtual resources the tasks are loaded to the suitable virtual machine according to the cost matrix table. If the no. of incoming tasks is not equal to the no. of existed virtual resources then we employ the assignment mechanism algorithm to detect the virtual machine for the requested application and implement the round robin technique to perform the tasks which are available in the queue depending on mentioned time slots of the respective virtual machines.

V. Proposed Load Balancing Technique

In the proposed mobile virtual cloud network will be form and will make available to all nodes for execution but it is very critical to find the optimize node for the execution in the given network so that we are using bidding resource allocation algorithm. This will work like to the bidder. In this way we can form the virtual cloud network with total cloud implementation.

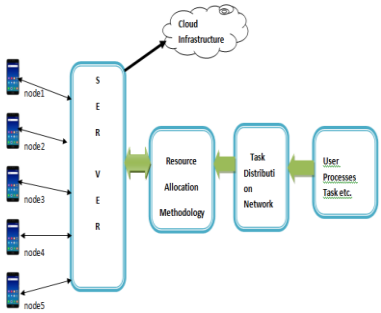

Figure 2: Proposed Architecture

The implementation is done on the resource availability of available resources in MCC. whenever the different node are available in network and to execute the certain task then finding out the efficient and optimized node is too much important in this we are implementing the architecture using agent based search mechanism which effectively help us in load balancing of in the allocation of the various task and manipulating the optimal resources

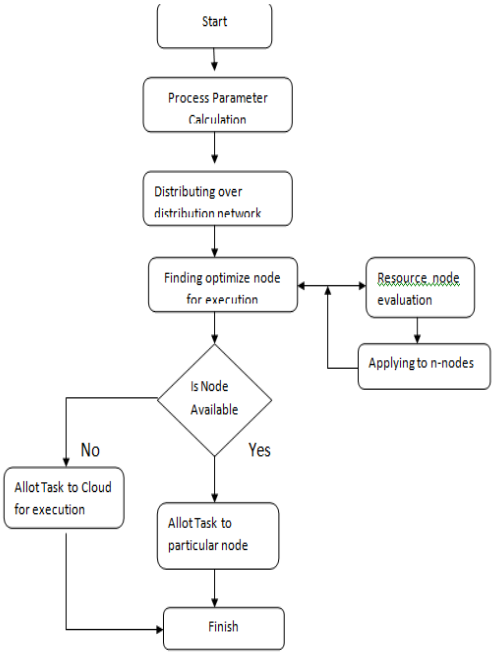

Figure 3: Proposed flowchart

In our proposed system the existing resources of cloud are taken in consideration for giving high performance to mobile and other devices which is cost consuming. So that it is possible to access the resources of joined node in the network in which the execution can be distributed to the another node and the device can be optimized. So that we are proposed the effective load balancing mechanism in support of the cloud in which the interconnected

node can form the virtual cloud network. In the proposed mobile virtual cloud network will be form and will make available to all nodes for execution but it is very critical to find the optimize node for the execution in the given network so that we are using bidding resource allocation algorithm. This will work like to the bidder. In this way we can form the virtual cloud network with total mobile cloud implementation.

VI. Conclusion

This paper presents implementation formula which can resolve the Virtual machine programming management at a lower place the dynamic atmosphere of the amount of VMs and requests on Cloud computing. We have projected a flow chart for load balancing in cloud computing environments supported behavior of honey bee forage strategy. The tasks are to be send to the under loaded machine and like forage bee consecutive tasks also are sent there to Virtual Machine until the machine gets overloaded as flower patches exploitation is completed by scout bees. Honey bee behavior galvanized load balancing, improves the general turnout of process and priority based balancing focuses on reducing the makespan , time a task must help a queue of the VM. Thus, it reduces the response of your time of VMs. The experimental results show that the formula is effective when put next with existing algorithms. Our approach illustrates that there's a big improvement in average execution time and reduction in waiting time of tasks on queue. Results show that our formula stands smart while not increasing further overheads.

Future Work

In future, there's scope for improvement within the algorithms. We have a tendency to arrange to improve this algorithm by considering alternative QoS factors of tasks. The performance of the given algorithms can even be augmented by variable totally different parameters. Projected algorithm remains a promising and fascinating algorithm, which might still be extensively employed by researchers across various fields. Its potential advantage of being simply hybridized with totally different metaheuristic algorithms and parts makes it robustly viable for continuing utilization for additional exploration and improvement prospects in more years to return.

References

[1] K Pathak, G Vahinde, “Comparison Of Particle Swarm Optimization And Genetic Algorithm For Load Balancing In Cloud Computing Environment”, International Journal of Research in Computer & Information Technology (IJRCIT) Vol. 1, Issue 1, 2015, ISSN: 2455-3743,

[2] AnoopYadav, “Comparative Analysis of Load Balancing Algorithms in Cloud Computing”, International Journal of Enhanced Research in Management & Computer Applications ISSN: 2319- 7471, Vol. 4 Issue 9, September-2015

[3] Yongfei Zhu, Di Zhao, Wei Wang, and Haiwu He, “A Novel Load Balancing Algorithm Based on Improved Particle Swarm Optimization in Cloud Computing Environment”, Springer International Publishing Switzerland 2016, HCC 2016, LNCS 9567, pp. 634–645, 2016, DOI: 10.1007/978-3-319-31854-7_57

[4] GeethaMegharaj, Dr. Mohan K.G, “A Survey on Load Balancing Techniques in Cloud Computing”, IOSR Journal of Computer Engineering (IOSR-JCE) e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 18, Issue 2, Ver. I (Mar-Apr. 2016), PP 55-61, DOI: 10.9790/0661-18215561,

[5] RenGao and Juebo Wu, “Dynamic Load Balancing Strategy for Cloud Computing with Ant Colony Optimization”, Future Internet 2015, 7, 465-483; doi:10.3390/fi7040465, ISSN: 1999-5903,

[6] Anju Baby, “Load Balancing In Cloud Computing Environment Using Pso Algorithm”, International Journal for Research in Applied Science And Engineering Technology (IJRASET), Vol. 2 Issue IV, April 2014, ISSN: 2321-9653,

[7] ElinaPacini, Cristian Mateos, and Carlos GarcíaGarino, “Dynamic Scheduling based on Particle Swarm Optimization for Cloud-based Scientific Experiments”, HPCLatAm 2013, Session: Evolutionary Computation & Scheduling, Mendoza, Argentina, July 29-30, 2013,

[8] FahimehRamezani· Jie Lu, Farookh Khadeer Hussain, “Task-Based System Load Balancing in Cloud Computing Using Particle Swarm Optimization”, International Journal of Parallel Programming, DOI 10.1007/s10766-013-0275-4,

[9] C. Clark, K. Fraser, S. Hand, J. G. Hansen, E. Jul, C. Limpach, I. Pratt and A. Warfield: "Live Migration of Virtual Machines", Proceedings of the 2nd USENIX Symposium on Networked Systems Design and Implementation, 2005.

[10] F. Travostino, P. Daspit, L. Gommans, C. Jog, J. Mambretti, I. Monga, B. Oudenaarde, S. Raghunath and P. Wang: "Seamless Live Migration of Virtual Machines over the MAN/WAN", Elsevier Future Generation Computer Systems, Vol.22, 2006.

Authors Profile

Chalanati Durga Venkata Peddintla Kumari presently pursuing M.Tech (CSE) degree from Kakinada Institute of Engineering & Technology Corangi, KKD-Yanam Affiliated to JNTU Kakinada. Her Area of interests include Computer Networks, Network Security, Cryptography, Cloud Computing and Software Testing.

Mr. S.V.Krishna Reddy, excellent teacher, Received M.Tech (CSE) from Kakinada Institute of Engineering and Technology, affiliated to JNT University, Kakinada is working as Assistant Professor, Department of Computer science engineering, Kakinada Institute of Engineering and Technology. He has 8 years of teaching experience in various engineering colleges. His area of Interest includes Hadoop and Big Data, Artificial Intelligence and other advances in computer Applications.

