iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 6 **Issue:** I **Month of publication:** January 2018

DOI: http://doi.org/10.22214/ijraset.2018.1014

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor :6.887
Volume 6 Issue I, January 2018- Available at www.ijraset.com

Load Balancing Techniques – A Review

Kuldeep[1], Jai Bhagwan[2]
[1,2] *Department of Computer Science Engineering, Guru Jambheshwar University of Science & Technology, Hisar*

Abstract: ***In recent years, cloud has been recognized as one of the emerging worldwide technologies in which shared resources are provided on the user's request at any time. The whole work of cloud is based on three service models (IaaS, SaaS and PaaS) and four deployment models (Public, Private, Community and Hybrid). In the cloud computing, Load balancing is one of the most challenging issues. It is a mechanism that disseminates the dynamic workload equitably over every node in the entire cloud to avoid a circumstance where few nodes are intensely loaded while others are idle or doing little work. It accomplishes a high client fulfilment and resource usage proportion, thus enhancing the overall performance and resource utilization of the system. In this paper, we describe some of the basic load balancing algorithms and literature review on the various algorithms proposed by different authors.***

Keywords: Cloud Computing, Load Balancing, Round Robin, FCFS, Throttled.

I. INTRODUCTION

Cloud computing is an advancing range that enables clients to sort out applications with improved versatility, accessibility and fault tolerance. Cloud computing gives a web-based platform that is utilized for computer technology. It describes a decent variety of processing ideas [1]. Cloud computing collects all the computing resources and handles them naturally. Nowadays world relies on cloud computing to store the public as well as individual data [2]. Cloud computing gives applicable hardware, software and service according to the requirement that clients set forward. A cloud computing structure is arranged by its on-require self-service, access over the web, pooling of resources, the versatility of service accessibility and estimation of services used by singular clients [2][3][8]. Cloud computing gives an aggregate gathering of resources, including information storage space, computer processing power networks and concentrated corporate and client application [3].

In cloud computing there are four deployment models [2][5]:

1) Private Cloud - The type of cloud, which is used by one authorization (single person or organization).
2) Public Cloud - The type of cloud is used publically (group of persons or organization).
3) Hybrid Cloud - It is the combination of public and private cloud.
4) Community Cloud - A shared cloud is created by two or more than two organization for sharing information between each other.

Three service models on which all the work of cloud is done are below [4][5]

IaaS - Infrastructure as a Service is the basic service model. In IaaS, service providers mainly provides the hardware related services like processing device, network, storage, memory and other fundamental devices. Some famous IaaS service providers are Amazon EC2, Amazon S3, Flexi scale and Rac kspace Cloud Servers. PaaS - It allows a user to develop an application and deploy in the cloud by providing a platform. Main service providers of development tool are Google's Application Engine, Salesforce.com, and Microsoft Azure. SaaS - The capability provided to the consumer is to use the provider's applications running on a cloud infrastructure. Some of the SaaS service providers are Salesforce.com, online Google, Yahoo, Hotmail and many others.

II. LOAD BALANCING

In cloud computing, a CSP (Cloud Service Provider) uses load balancing for distribution of resources across all the server, datacenters and hard drives used in the network system of cloud. It allots all application requirements across any number of application distributions nodes which are positioned in different datacenters [3][7]. Load balancing can be generally characterized by numerous techniques. These are: Integrated or decentralized, active or static, and periodic or non-periodic [7].

Load balancing is the way toward enhancing the performance of the system by moving of workload among the processors. Workload of a machine implies the combined processing time it requires to execute every task assigned to the machine [1][6]. Virtual machine's balance consistently implies that any of the accessible machines is not idle or partially loaded while others are having high load [7]. The advantages of circulating the workload include expanded resource utilization ratio which additionally improving the general execution and overall performance thereby achieving extreme customer fulfilment. In cloud computing, if clients are expanding, the load will likewise be expanded, the expansion in quantity clients will prompt poor execution as far as resource usage if the cloud supplier isn't arranged with any good system for load balancing and furthermore the limit of cloud

servers would not be used appropriately [8][9]. This will hijack or quizzes the execution of substantially loaded node. On the off chance some good load balancing technique is implemented, it will similarly isolate the load and along these we can maximize resource utilization. One of the cloud computing major issues is to separate the workload dynamically [10].

In cloud computing environment, a few resources are intensively loaded while others are idle or doing little work, its fundamental reason is the random arrival of a task and random usage of CPU service time, thus in cloud computing load balance and resource control are principle key challenging issues. To achieve minimum response time, less energy efficiency, optimal resource utilization and remove overload of a node it provides a procedure to allocate workload across multiple computers or nodes. The fundamental reason behind load balancing is to enhance execution performance by balancing the load among different resources viz. network links, disk drives, and central processing units. Different load balancing algorithms are used to distribute the load to different systems. Two type of load balancing are:

A. Static load balancing - In the static load balancing algorithms, the load distribution does not depend upon the present condition of the cloud system. The information about the resources of the system and applications is needed [1]. While arrival of jobs occurs, the execution of the virtual machines is determined. As per the execution requirement, the master processor allocates the workload to slave processors. The allocated work is as a result performed by the slave processors and the outcome has come back to the master processors [3][6][10].

B. Dynamic Load Balancing - In dynamic type of load balancing algorithms, it is to settle down load balancing by using present condition of availability of resources. In this way, the moving of the load is relied upon the present condition of the system [1]. Here, the processes are picked up to move from an overused machine to underused machine dynamically for getting better performance [6][10].

III.BASIC ALGORITHM OF LOAD BALANCING

A. The basic algorithm of load balancing is described below [1][6][9][23]

1) Round Robin Load Balancing: Round Robin (RR) algorithm focuses on the fairness. It is a static load balancing technique in which tasks are scheduled on the basis of time quantum. RR uses the ring as its queue to store jobs. The fundamental preferred standpoint is that it is a starvation free technique. Each procedure will be executed by CPU for settled time cut. So along these lines no procedure left for its round to be executed by the CPU. All tasks will performed with no prioritization. It chooses the load randomly and leads to the circumstance where a few nodes are intensely loaded and few are softly loaded. In spite of the fact that the algorithm is exceptionally straightforward and simple to actualize however there is an extra load on the scheduler to choose the time quantum.

2) Min–Min algorithm: At first, the Min-Min algorithm is used to find the minimum expected time of all tasks in meta-task. The task having the minimum expected completion time is selected and assigned to the corresponding resource. This process is continued iterated until meta-task is not empty. Here, a task having bigger execution time has to wait for the completion of smaller tasks.

3) Max-Min algorithm: The Max-Min algorithm is used to calculate the completion time of each task as per the available resources. A task which has overall maximum completion time is scheduled over a resource with overall minimum execution time. This step is repeated until meta-task is not empty. Here, the waiting time of a larger task is reduced.

IV.LITERATURE REVIEW

Many techniques have been proposed for load balancing in the cloud and some of them are explained below.

1) K. A. Nuaimi et al. (2012) described the most known contributions in the literature for load balancing in cloud computing. The authors classified the load balancing algorithms into two types: static algorithms and dynamic algorithms. Static load-balancing algorithms have been developed for cloud computing based on the ability of the node to process new requests and assigned tasks. The process is based on prior knowledge of the nodes properties and capabilities. These would include the node's processing power, memory and storage capacity, and most recent known communication performance [11].

2) Sethi et al. (2012) proposed a new load balancing approach using fuzzy logic with RR (round robin) algorithm. The approach is based on number of parameters that are processor speed, load on virtual machines, response time etc. This approach maintains the current information of every VM and all the requests allotted to VMs. When a process request is arrived, the load balancer examine all the VMs for least loaded assigned to it, rather than if there are more than one VMs present, assignment of

the process is checked with the fuzzy logic on the basis of processor speed and load. The result shown better performance than RR algorithm [12].

3) *Sharma et al. (2013)* proposed a new approach for load balancing to improve response time of every Virtual Machine (VM). This approach gathered data about all VMs in a list and utilized it for allotment of virtual machine whether the status is idle or not. At the point when another demand is received, load balancer looked into the table and recognized the machines having low load. The acknowledge comes back to the data enter and allot the resources to the processes. When the process is completed, it tells to the datacenter and de-allocate the resource. The main disadvantage of this approach is wating queue length is bigger in the case of bigger workload and allotment of process depend upon only available status of machines [13].
4) *Buyya Rajkumar et al. (2014)* iintroduced an architecture Software-Define-Clouds (SDCs) focusing on a variety of applications including computer, mobile, data intensive application in web and enterprise environments. Here, the authors set different elements that comprise the architecture and evaluated the potential of SDCs in two use cases QoS-aware Bandwidth and Bandwidth-aware allocation, energy-efficient Virtual Machine placement through simulation. The open challenges and opportunities arising from this emerging area were also discussed [14].
5) *A. Rodriguez Maria et al. (2014)* proposed a method for combining resource provisioning and scheduling strategy for execution of scientific workflows on IaaS clouds. The main objective of the scenario was to reduce the overall execution time and cast within the user defined deadline. The problem was solved using the meta-heuristic optimal algorithm say Particle Swarm Optimization (PSO). The experimental results shown that the proposed solution provided better performance as compared to the state-of-the-art algorithms SCS and IC-PCP. In most cases, IC-PCP was failing to meet the application deadline but introduced an approach meets. Defining an approach to meet deadlines as SCS and IC-PCP met the deadline. The authors produced schedules with lower execution cost [15].
6) *Singh et al. (2014)* proposed an approach of heterogeneous load balance to circulate load among various servers. The authors created VMs of different datacenters having various RAM sizes, size of secondary memory, processor speed etc. After getting requests from client, the allotment of available resources or VM was done on the basis of RAM/memory size and VM allocation in the list was changed from available to busy mode. After process request completion, the VM allocation in the list was changed to the available. Disadvantage of this approach is allocation of VM which is having high RAM rather ignoring the other criteria such as processing speed and others [16].
7) *Rathore Monika et al. (2015)* introduced a technique based on honey bee forage strategy. In this mechanism, the tasks were sent to the under-loaded machine and continue to send the task to VM to VM; after that the VMs set to overloaded. Honey Bee Behaviour galvanizes balanced the load and minimized the makespan, reduce the waiting time because of priority based balancing, it also reduced the response time of VMs. The result of the approach shown average execution time and waiting time to rejoice [17].
8) *Abhay Kumar Agarwal et al. (2015)* proposed a new static load balancing technique for cloud computing. The proposed technique used both the approach active monitoring and throttle load balancing. In this paper, the overall result on the response time and processing time were better than the throttle load balancing algorithm [18].
9) *B. Nagpure Mahesh et al. (2015)* proposed a technique that used for dynamic resource allocation system. This technique decreased the overload in serves effectively by allocating resource evenly among VMs. Here the condition was that the capacity of a physical machine would be sufficient. The proposed work achieved optimizes performance in term of server resource utilization and it reduced energy consumption and task migration among VMs [19].
10) *Luo Jianying et al.(2015)* introduced an approach to study how to leverage both the geographic and temporal variation of energy variation to reduce energy cost for distributed IDCs. Authors proposed a novel architecture and two algorithms for united spatial and temporal load balancing. Result Analysis shown that the algorithms have a low computational complexity, require a relaxed accuracy in electricity price estimation and guarantee a service completion time for user requests [20].
11) *R. Ashalatha et al. (2015)* proposed an algorithm that used a dynamic load balancer which provided higher user satisfaction and maximum resource utilization. Dynamic load balancer achieved high performance and greater resource utilization of the system. Dynamic allocation of application made the sessions to be transferred from one side to another as needed. A dynamic load balancer distributed the scalable workload among all the nodes in cloud computing. DLB used to create new instances. The resources were allocated when required and released when no need of resources, as a result it reduced the cost of resource allocation [21].
12) *Vig Aarti et al. (2015)* studied this paper that many load balancing techniques had been introduced having some advantages and disadvantages. There are mainly two types of load balancing technique - static and dynamic. Static balancer has

disadvantages of monitoring and simulation. It has the limitation of designing a model having heterogeneous nature of clouds. On the other end, the dynamic balancer has advantages as it models the heterogeneous nature of cloud easily with a limitation as it is very difficult to simulate it [22].

13) *B Pavithra et al. (2016)* presented a comparative analysis in terms of performance among various load balancing techniques like WFCFS, RR and TLB. Optimal one could be found based on metrics like energy consumption, execution and response time, processing costs, and then found out optimal solution that provides better solution data center, in term minimizing energy consumption processing cost, execution and response time of the VMs. The simulation was performed using Cloud sim and Cloud Analyst tools to obtain the result. Results shown WFCFS algorithm provides better performance in all factors, better resource utilization and energy consumption. WFCFS is the best technique of energy consumption. [23].

TABLE 1. EXISTING ALGORITHMS COMPARISON

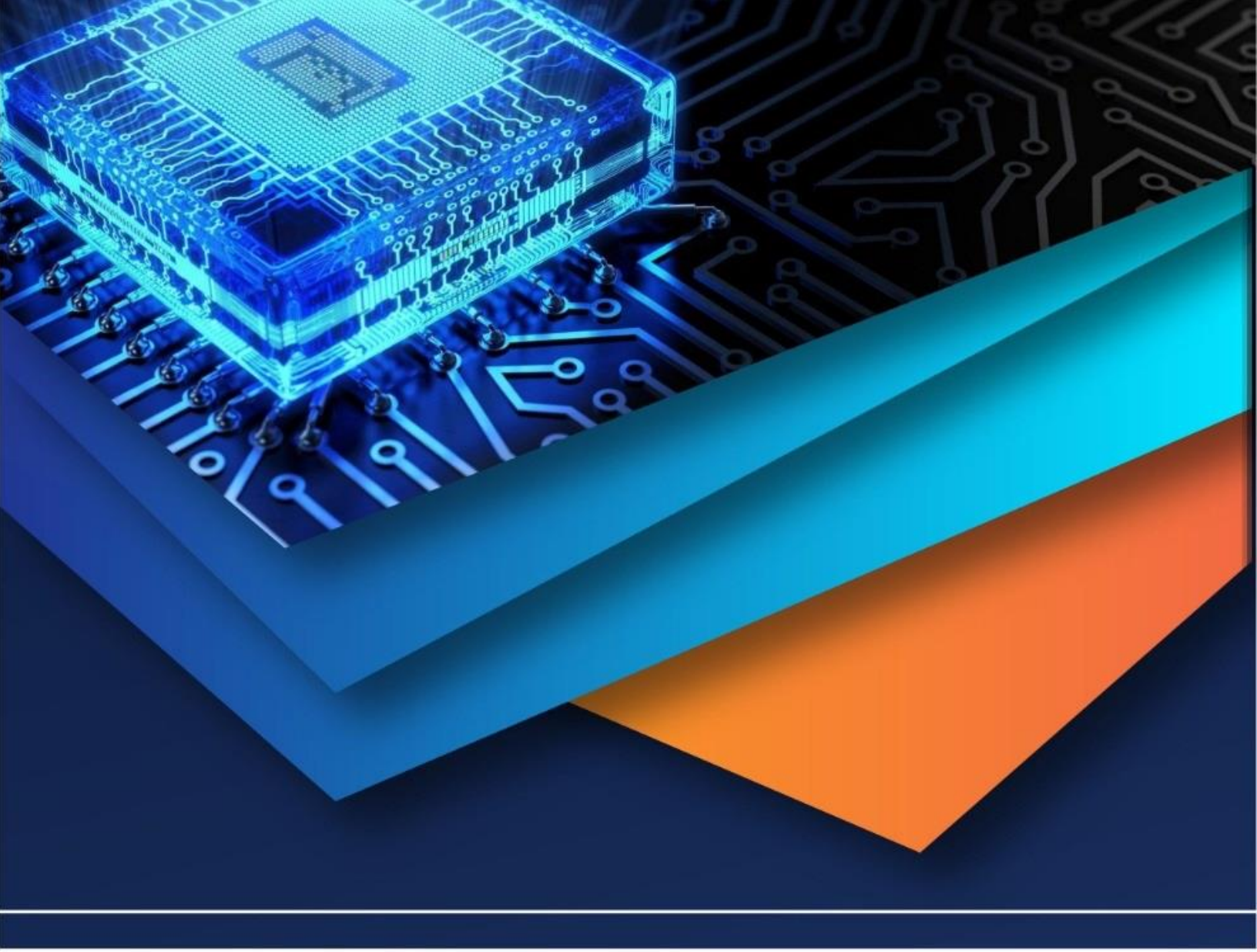
Algorithm	Simulation Parameters	Findings	Environment
Round Robin[9]	Completion Time	Better Than FCFS	Cloud Analyst, Cloudsim
Min-Min[9]	Execution Time	Better Than RR, FCFS	Workflowsim
Max-Min[9]	Execution Time	Better Than RR, FCFS	Workflowsim
Round Robin Load Balancing With Fuzzy Logic[12]	MIPS, Memory Capacity , Waiting Time	Better Than RR	Cloudsim
The efficient and enhanced algorithm[13]	Response time	-	Cloudanalyst
ANT Colony Optimization[14]	Completion Time	Perform Better Than Particle Swarm Optimization, GA	Cloudsim
Heterogeneous Load Balancing Algorithm[16]	MIPS, Memory, Storage	Not Mentioned	Not Mentioned
Honey Bee Galvanized Load Balancing Formula (HBBLB)[17]	Response time, makespan	-	Cloudsim
Static Load Balancing Algorithm[18]	Overall Response Time And MIPS	Perform Better Than Throttled Load Balancing Algorithm	Cloudanalyst

V. CONCLUSIONS

Load balancing is the most challenging issue in cloud computing. In this paper, we have discussed various load balancing techniques. A survey has done on various new approaches such as dynamic load balancer (DLB), RR with fuzzy logic, ACO, Static Load Balancing etc. All these methods have been compared on the basis of response time, completion time, MIPS etc. Due to rapid growth in user numbers day by day, a better approach is required to minimize the makespan, execution time by adjustment of the VMs load and provides better cost optimization solution for the client.

REFERENCES

[1] Akanksha Mathur, Virender Singh Shekhawat, "Load Balancing in Cloud Computing: A Review", International Journal of Innovative Research in Computer and Communication Engineering, Vol. 3, Issue 11, November 2015.
[2] M. Ahmed, M. S. Abu, R. Chowdhury, A. Mustaq and M. M. Hasan Rafee, "An advanced survey on cloud computing and state-of-the-art research issues,"International Journal of Computer Science Issues, Vol. 9, Issue 1, pp. 0814-1694, 2012.
[3] Mohit et al., "A Survey of Existing Load Balancing Algorithms in Cloud Computing ", International Journal of Advanced Research in Computer Science and Software Engineering 6(4), pp. 840-844, April- 2016.
[4] R. K. Ranjan and D. Rani, " A Comparative Study of SaaS, PaaS and IaaS in Cloud Computing," International Journal of Advanced Research in Computer Science and Software Engineering, Vol. 4, issue 6 , June 2014.
[5] B. Sahu and R. Tiwari, "A Comprehensive Study on Cloud Computing," International Journal of Advanced Research in Computer Science and Software Engineering, Volume 2, Issue 9,pp.33-37, September 2012.
[6] D. Saranya, L. Sankara Maheswari, " Load Balancing Algorithms in Cloud Computing: A Review", International Journal of Advanced Research in Computer Science and Software Engineering 5(7), pp. 1107-1111, July- 2015.
[7] Sukhvir Kaur, Supriya Kinger, " Review on Load Balancing Techniques in Cloud Computing Environment", International Journal of Science and Research (IJSR) ISSN (Online): 2319-7064, 2014.
[8] Tushar Desai, Jignesh Prajapati, "A Survey Of Various Load Balancing Techniques And Challenges In Cloud Computing",International Journal Of Scientific & Technology Research Volume 2, Issue 11, Issn 2277-8616, November 2013.
[9] Dharmesh Kashyap, Jaydeep Viradiya, "A Survey Of Various Load Balancing Algorithms In Cloud Computing", International Journal of Scientific & Technology Research Volume 3, Issue 11, November 2014.
[10] F. Kherani and Prof. J. Vania "Load Balancing in cloud computing," International Journal of Engineering Development and Research, Vol. 2, issue 1, 2014.
[11] Klaithem Al Nuaimi, Nader Mohamed, Mariam Al Nuaimi and Jameela Al-Jaroodi," A Survey of Load Balancing in Cloud Computing: Challenges and Algorithms," IEEE Second Symposium on Network Cloud Computing and Applications, IEEE, 2012.
[12] Sethi, S., Anupama, S., and Jena, K., S, "Efficient load Balancing in Cloud Computing using Fuzzy Logic", IOSR Journal of Engineering (IOSRJEN), 2(7): pp. PP 65-71, 2012.
[13] Sharma, T. and Banga, V. K., "Proposed Efficient and Enhanced Algorithm in Cloud Computing", International Journal of Engineering Research & Technology (IJERT), 2(2), 2013.
[14] R. Buyya, R.N. Calhetros, J. Son, A. V. Dastjerdi and Y. Yoon, "Software-Defined Cloud Computing: Architectural Elements and Open Challenges," International Conference on Advances in Computing, Communication and Informatics (ICACCI), IEEE, 2014.
[15] M. A. Rodriguez and R. Buyya, "Deadline Based Resource Provisioningand Scheduling Algorithm for Scientific Workflows on Clouds," 3rd International Symposium on Computing and Networking, vol. 2, no. 2, pp. 222–235, IEEE, Apr. 2014.
[16] Singh, A., Bedi, R., and Gupta, S., "Design and implementation of an Efficient Scheduling algorithm for load balancing in Cloud Computing", International Journal of Emerging Trends & Technology in Computer Science (IJETTCS), 3(1), 2014.
[17] M. Rathore, S. Rai and N. Saluja, "Load Balancing of Virtual Machine Using Honey Bee Galvanizing Algorithm in Cloud", International Journal of Computer Science and Information Technologies, vol. 6(4), IJCSIT, 2015.
[18] Abhay Kumar Agarwal, Atul Raj, "A New Static Load Balancing Algorithm in Cloud Computing", International Journal of Computer Applications (0975 – 8887) Volume 132 – No. 2, December 2015.
[19] M. B. Nagpure, P. Dahiwale, and P. Marbate, "An Efficient Dynamic Resource Allocation Strategy for VM Environment in Cloud," International Conference on Pervasive Computing(ICPC), IEEE, 2015.
[20] J. Luo, L. Rao, and X. Liu, "Spatio-Temporal Load Balancing for Energy Cost Optimization in Distributed Internet Data Centers," IEEE Transaction on Cloud Computing, vol. 3, no. 3, pp. 387–397, IEEE, Jul. 2015.
[21] R. Ashalatha and J. Agarkhed, "Dynamic load balancing methods for resource optimization in cloud computing environment," in India Conference (INDICON), pp. 1–6, IEEE, 2015.
[22] A. Vig, R. S. Kushwah, and S. S. Kushwah, "An Efficient Distributed Approach for Load Balancing in Cloud Computing," 2015 International Conference on Computational Intelligence and Communication Networks (CICN), pp. 751–755,IEEE, 2015.
[23] B. Pavithra and R. Ranjana, "A comparative study on performance of energy efficient load balancing techniques in cloud," International Conference on in Wireless Communications, Signal Processing and Networking (WiSPNET), pp. 1192–1196, IEEE, 2016.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY