International Journal of Latest Trends in Engineering and Technology
Special Issue SACAIM 2016, pp. 309-316
e-ISSN:2278-621X

SCHEDULING AND LOAD BALANCING TECHNIQUES IN CLOUD COMPUTING: A SURVEY

Mithun Dsouza[1], Mohammed Rizwan[2], Ramnath Gaonkar[3] and Dr. S. Sathyanarayana[4]

Abstract—Cloud computing has become popular due to its attractive features. The load on the cloud is increasing tremendously with the development of new applications. Load balancing is an important part of cloud computing environment which ensures that all devices or processors perform same amount of work in equal amount of time. In this paper we are mentioned about different techniques in load, we aim to provide a structured and comprehensive overview of the research on load balancing algorithms in cloud computing. This paper surveys the state of the art load balancing tools and techniques over the period of 2004-2016.

Keywords: cloud computing, load balancing, load balancing algorithms.

I. INTRODUCTION

Cloud computing is the one of the emerging technology in the world and many companies are trying to do the work faster and faster in the competitive world. Cloud computing has been adopted by organization which includes, social networking websites, online applications are designed by Google app managers and by Google doc which are some of the important implementation and a step ahead in cloud computing. Cloud computing accumulates all the computing resources and maintains them automatically. Today's generation depends on cloud computing to store the public as well as personal information. Cloud computing supplies relevant hardware, software and services according to the users need. Cloud computing is a new technology and it providing online resources and online storage to the user's .It provide all the data at a lower cost. In cloud computing every users can access resources all the time through internet. They need to pay only for some of the resources as much they use .In Cloud computing cloud provider outsourced all the resources to their client. The main problem is load balancing in cloud computing. Load balancing helps to distribute all loads between all the nodes [2].

Cloud computing provides flexible way to maintain data and files, which involves virtualization, distributed computing, and web services. It also has several elements like client and distributed servers. The main target of cloud computing is to provide maximum number of services with minimum cost at any point of time. Nowadays, there are more than hundred millions of devices

[1] *AIMIT, St Aloysius College Mangalore, Karnataka, India*
[2] *AIMIT, St Aloysius College Mangalore, Karnataka, India*
[3] *AIMIT, St Aloysius College Mangalore, Karnataka, India*
[4] *First Grade Women's College, Mysore, Karnataka, India*

connected to the Internet. These devices submit their request and receive the response without any delay [3]. Figure 1, it shows different devices (mobiles, PCs, laptops) connect and access the data from a cloud at any given time.

Figure 1. A Cloud Computing Scenario

II. LOAD BALANCING IN CLOUD COMPUTING

Load Balancing strives to distribute the load in equal proportions across resources depending on recourse capacity so that each resource is not overload or underutilized in a cloud system. Moreover, it is considered to be a NP-hard optimization problem that has 2 types: dynamic load balancing and Static load balancing. Dynamic load balancing is to rebalance the system while running when detecting overloaded VMs while static load balancing is to balance the system from the start by scheduling. Static load balancing is preferable as it avoids VM migration costs and provides better execution time and quality of service (QoS)[4].

There are several characteristics of load balancing such as: equal division of work across all the nodes, facilitation in achieving user satisfaction, improve the system performance, reduce response time, and provide services to achieve complete resource utilization [3]. Figure 2 shows the load balancing in cloud computing. As an example, if we make one application on cloud and thousands of users are expected to access it at any point of time. Therefore, response time to thousandnumber of people will be too slow and servers will become busy very quickly, resulting in slow response and unsatisfactory users. By applying load balancing on our application it become faster, then work will be distributed at other nodes and we can get high performance and better response [5].

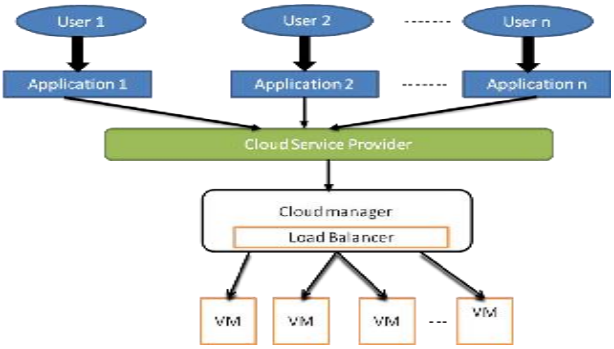

Figure 2. Load balancing in cloud computing

III.LITERATURE SURVEY

In load balancing algorithm finding the effort of balancing the server workloads[1]. Load Balancing is one of the main issues in cloud computing, as it improves the system response time, power consumption and utilizes resources efficiently [4]. The process in which the load is divided among several nodes of distributed system is called load balancing in cloud computing. Load balancing assists the cloud computing through algorithms [3]. Many works has been done to balance the load in order to improve the better performance and avoid over utilization of

resources.Different load balancing algorithms are discussed inRR(round robin),RA(random allocation), Min-Min, Max-Min and other existing algorithm's. Load balancing algorithms are classified into two types: Static load balancing algorithms and Dynamic load balancing algorithms [5].

A. Static Load Balancing Algorithm-

R. lee [16] proposed a static load balancing algorithm in cloud load balancing. In Static algorithm the Scheduling assignment of tasks to processors is done before program execution begins i.e. in run time. Scheduling algorithms decision is based on information about task execution times, processing resources etc. which is assumed to be known at compile time [3]. In this section, we provide a detailed discussion on the existing load balancing algorithms for cloud. Static load balancing algorithms are not preemptive, therefore each machine has at least one task assigned for itself [6]. Static algorithms generally do not consider dynamic changes of these attributes at executing time. In addition, these algorithms cannot adapt to load changes during run-time.

a.. Round Robin Load Balancer (RR)-

Radojevic proposed an algorithm called CLBDM (Central Load Balancing Decision Model). CLBDM is an improvement method in Round Robin Algorithm which is based on session switching at the application layer. A Round Robin is an arrangement of choosing all elements in a group equally in some well advised order, usually from the top to the bottom of a list and then again it'sstartingat the top of the list and so on. In this algorithm all the processes are divided between all processors. In this each process is allocated to the processor in a round robin order. The work load distributions between processors are equal [2]. .Different process are not having the same process time. At sometimes the some nodes are heavily loaded and others stand inactivein web servers where http requests are of similar nature and distributed equally then RR algorithm is used.

b.Min-Min Scheduling Algorithm-

H.Chen [18] proposed a Min-Min algorithm in load balancing. In Min-Min the smallest task is assigned to the fastest resource. Here the task is removed from the set and same process is repeated and this method is simple.Here smallest execution time value. In presence of more small tasks, it shows best result.It starts with a set of all unassigned tasks .In this minimum execution time for all tasks is found. Then after that among these minimum times the minimum value is selected. After task with minimum time schedule on machines. After that the execution time for all other tasks is updated on that system then again the same procedure is followed until all the tasks are assigned on the resources. The main problem of this algorithm is has a starvation [2].

c. Max-Min Scheduling Algorithm-

U.bhoi [19] proposed a Min-Min algorithm in cloud load balancing. Max-Min algorithm is almost same as Min-Min algorithm, but in Max-Min it chooses the task with maximum value and gives into respective machine.In this algorithm first finding out minimum completion times, then the maximum value is selected which is the maximum time among all the tasks on any resources. After that maximum time finding, the task is allocated on the particular selected machine. Then the execution time for all tasks is upgraded on that machine, this is done by adding the execution time of all assigned task to the

execution times of other tasks on that machine. Then all assigned task is removed from the list that executed by the system [2].

B. Dynamic Load Balancing Algorithm-

N. Haryani [21] proposed a Dynamic load balancing algorithm in cloud load balancing. Dynamic load balancing algorithm which finds for the lightest server in the network and then designated appropriate load on it. In this, work load is distributed among the processors at runtime. Here the algorithms in this category are considered complex and it's having better fault tolerance and overall performance [9].Dynamic algorithms are well suited in cloud computing environment because they distribute work at run time and assign suitable weights to the servers. A lightest weight server is search in network and preferred by this algorithm [3]. In dynamic load balancing the load distributes among the nodes throughout run-time. If load balancer finds high usage of CPU the request is send to the next node to handle the load, current state of the system is used [3]. Decisions on LB are based on current state of the system. No prior knowledge is compulsory. Thus it is better than static approach [10].

a. Ant colony algorithm-

Marco Dorigofirst estrablished the Ant System (AS) in his Ph.D thesis in 1992. Today it is one of the best optimization technique, here which finds the shortest path. The deposition of pheromone and the ant moves approximately at the same speed and at the same rate. And that pheromone attracts the remaining ants to move on same path. So, more ants move on same path have higher concentration of pheromone and the evaporation rate is too low on shorter path, that's why ants chooses the shorter path [1]. Different ant colony algorithms also introduce to balance the load applying ant behavior for searching food. Larger weight means that resource has high computation power. Load balancing ant colony optimization (LBACO) not only balance the load but also minimizes make span. All tasks are assumed to be mutually independent and computationally intensive [3]. In Ant Colony Faster information can be collected by the ants and minimizes make span.J. Kaur proposed an Ant Colony Algorithm in cloud load balancing[22].

b.Honey bee foraging algorithm-

Dhinesh et al proposed an algorithm after detail analysis of foraging behavior of honey bees [3]. Honey bee foraging algorithm it increases the throughput and it minimize the response time.Honey bee foraging algorithm is a decentralized honeybee-based nature-inspired load balancing technique for self-organization. It achieves global load balancing through local server action. This algorithm is derived from the behavior of honey bees for foraging and harvesting food [5]. Here it is an under loaded VM assigns a task, it updates number of priority tasks and load of VM to other tasks in waiting list. This approach helps other processes to choose their VM. If a task has high priority, then it selects a VM having minimum number of priority tasks. It does not take into consideration only load balancing but also keeps track of priorities of tasks which currently removed from heavy loaded machines [3].Network is over headed so search takes long time and no clarity about the number of ants.

c.Throttled load balancing-

M. Randless[23] proposed a Throttled load balancing in cloud load balancing. Throttled load balancing mainly used to manage the tasks.The Throttled Load Balancer (TLB) maintains a record of the state of each virtual machine and during allocation of a request the current load on the VM is not considered which can in turn increase the response time of a task [5].This algorithm depends upon the theory of suitable search of virtual machine. The task manager

makes a list of virtual machines. By using the list, client request allotted to the relevant machine. If the size and capability of the machine is suitable for request, then the job is given to that machine. This algorithm is better than round robin algorithm [3]. During allocation of a request the current load on the VM is not considered which can in turn increase the response time of a task [5].

d. OLB and LBMM-

Wang et al is proposed a combination opportunistic load balancing (OLB) and load balancing min-min (LBMM) algorithms to improve the performance of each tasks and it increases the task proficiency. All the tasks are given to each nodes in a specific manner and here the results are better than all other algorithms and also it is used in LBMM. In LBMM it starts by executing Min-Min algorithm at the first step. At the second step it chooses the smallest size task from the heaviest loaded resource and calculates the completion time for that task on all other resources. Then the minimum completion time of that task is compared with the makespan produced by Min-Min. If it is less than makespan then the task is reassigned to the resource that produce it, and the ready time of both resources are updated [5]. J. Uma [24] proposed an opportunistic load balancing in cloud load balancing.

IV.COMPARITIVE ANALYSIS OF SOME OF THE EXISTING ALGORITHMS

Table 1: Comparison of different Load balancing Algorithms

Scheduling algorithms	Title and Author	Advantages	disadvantages	Fairness	Complexity level	Overhead
Static load balancing algorithm [16]	Load Balancing Algorithms in Cloud Computing: A Survey of Modern Techniques By sidraaslam	At compile time decision about load balancing is made. The traffic is divided equally among the servers. Fewer complexes	There are only few load variations that are limited to the environment. It does not have ability to make changes at the end of run-time.	Yes	low	N/A
Round-Robin[17]	A Survey on Scheduling and Load Balancing Techniques in Cloud Computing Environment By Subhadra Bose Shaw	It is easy to understand, provides fairness and performs better for short CPU burst. There is fixed quantum time and also used priority.	Long time is taken for larger tasks. Context switches occurs more due to short quantum time and to achieve the high performance the job should be same. Longer jobs are keep on switching when quantum is slow.	Yes	low	high
Min-Min[18]	Load Balancing in Cloud System using Max Min and Min Min	In min-min response time is faster and also providing good performance.	Starvation. Machine and task variation cannot be predicted. Here it does not	no	low	high

	Algorithm By RajwinderKaur	In presence of more small tasks it shows the best result. This method is simple.	consider the existing load on a resourse.			
Max-Min[19]	Load Balancing in Cloud System using Max Min and Min Min Algorithm By RajwinderKaur	In max-min it reduces the makespan. Here requirements are prior known. So it works better.	Here the smaller jobs have to wait for long time. Here it takes long time to complete the task.	no	low	high
Dynamic Load Balancing Algorithm [21]	Load Balancing Algorithms in Cloud Computing: A Survey of Modern Techniques By sidraaslam	Here the tasks are distribute work at run time. Fault tolerance only check for the current state of system is required. Faster.	This method is complicated. Here constant check of each tasks. High complexity.	no	high	high
Ant colony algorithm [22]	Pros and Cons of Load Balancing Algorithms for Cloud Computing by urmilashrawankar	Here the best case scenario is provided that's why theunder loaded node is found at beginning of the search Decentralized, no single point of failure. Here ants can collect the Information faster.	Network overhead because of the large number of ants. Number of ants are not clear. Only availability of the node is considered. Nodes status change after ant's visits to them is not taken into account.	no	no	high
Honey bee foraging[3]	Load Balancing Algorithms in Cloud Computing: A Survey of Modern Techniques By Munamali shah	It works better under heterogeneous resources. This algorithm increases the throughput and minimizes the response time.	It increase in resources does not improve throughput equally. Here high priority tasks cannot works without virtual machine.	no	low	low
Throttled load balancing [23]	Load Balancing in Cloud Computing using Modified	TLB tries to distribute the load equally among the virtual machines.	It does not consider the present load on virtual machine. In this method the task is need to be	no	low	low

	Throttled Algorithm By Shridhar G. Domanal	Good performance and this method is used to manage the tasks.	waited.			
OLB+LBMM[24]	Pros and Cons of Load Balancing Algorithms for Cloud Computing by BhushanGhutke	It is simple. It improve the performance of each tasks. It increases the task proficiency.	Because of three layers this algorithm is slow than other algorithms.	no	high	low

V. CONCLUSION

The load balancing of the system is one of the greatest concern which should be worked and improved upon. Various techniques and algorithms are used to solve the problem. In this paper we survey various existing load balancing methods in different environments and have also mentioned their advantages, disadvantages and many other details which can be used by the researcher to develop more precise and efficient algorithms. A large number of parameters and different types of soft computing techniques can be included in the future for the better utilization and needs of the user. The various load balancing techniques are also being compared here which can be used for more improvements and research in future.

REFERENCES

[1] Ranjan Kumar Mondal, Enakshmi Nandi and DebabrataSarddar,"Load Balancing Scheduling with Shortest Load First" Vol. 8, No.4, (2015),pp.172.

[2] RajwinderKaur and PawanLuthra, "Load Balancing in Cloud Computing" SBS State Technical Campus/M.tech ,CSE,Student,Ferozepur, India, pages 374-375

[3] Sidra Aslam and manamalishah,"Load Balancing Algorithms in Cloud Computing: A Survey of Modern Techniques" in 2015 National Software Engineering Conference (NSEC 2015)pages 30-31

[4] Aya A. Salah Farrag,safiaabbas Mahmoud and EI Sayed M. EI-Horbaty "Intelligent Cloud Algorithms for Load Balancing problems: A Survey", 2015 IEEE Seventh International Conference on Intelligent Computing and Information Systems, pages 210-211

[5] Subhadra Bose Shaw and Dr. A.K. Singh , " A Survey on Scheduling and Load Balancing Techniques in Cloud Computing Environment" 2014-15 international conference on computer and communication technology, pages 88-89

[6] Hamid Shoja ,HosseinNahid and Reza Azizi ,"A COMPARATIVE SURVEY ON LOAD BALANCING ALGORITHMS IN CLOUD COMPUTING" IEEE – 33044 July 11- 13, 2014

[7] Klaithem Al Nuaimi, Nader Mohamed, Mariam Al Nuaimi and Jameela Al-Jaroodi,"A Survey of Load Balancing in Cloud Computing: Challenges and Algorithms" 2012 IEEE Second Symposium on Network Cloud Computing and Applications. Pages 138-139

[8] Dr. Thomas Yeboah, Prof. Odabi and Kamal Kant Hiran, "An Integration of Round Robin with Shortest Job First Algorithm for Cloud Computing Environment" International Conference On Management, Communication and Technology (ICMCT) April 2015, Vol III Issue –1 pages 2-3.

[9] Sushil Kumar and Deepak Singh Rana ,"Various Dynamic Load Balancing Algorithms in Cloud Environment: A Survey" International Journal of Computer Applications Volume 129 – No.6, November2015, pages 14-15

[10] Nikita Haryani,DhanammaJagli ,"Dynamic Method for Load Balancing in Cloud Computing" IOSR Journal of Computer Engineering (IOSR-JCE) e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 16, Issue 4, Ver. IV (Jul – Aug. 2014), PP 24-25

[11] Deepak B S, Shashikala S V and Radhika K R ,"Load Balancing Techniques in Cloud Computing: A Study" International Conference on Information and Communication Technologies (ICICT- 2014), pages 2-3

[12] SheetanshuRajoriya ,"Load Balancing Techniques in Cloud Computing: An Overview" International journal of science (IJSR) ISSN (Online): 2319-7064, pages 1619-1620

[13] KundanPagar and SachinPatil, "A Survey on Load Balancing Techniques in Cloud Computing"International Journal of Science and Research (IJSR) ISSN (Online): 2319-7064, pages 2252-2253

[14] Peter Mell, Timothy Grance, "The NIST Definition of Cloud Computing", NIST Special Publication 800-145, September 2011.

[15] google engine date, 10-november-2016

[16] R. Lee and B. Jeng, "Load-balancing tactics in cloud," Int. Conf. CyberEnabledDistrib. Comput.Knowl.Discov., pp. 447–454, Oct. 2011.

[17] D.Powar, S. SwaroopMoharana and R. D. Ramesh "analysis of load balancers in cloud computing," International Journal of Computer Science and Engineering (IJCSE), vol. 2, no. 2, pp. 101–108, 2013.

[18] H. Chen and F. Wang, "User-priority guided min-min scheduling algorithm for load balancing in cloud computing."Parallel Computing Technologies (PARCOMPTECH), National Conference IEEE, 2013

[19] U. Bhoi, P. N. Ramanuj, and W. S. Email, "Enhanced max-min task scheduling algorithm in cloud computing," International Journal of Computer and Information Technology, vol. 2, no. 4, pp. 259–264, 2013

[20] Ibrahim abdullahi, Process scheduling in LJF algorithm. A proposed framework for starvation problem, January 2012.

[21] N. Haryani and D. Jagli, "Dynamic method for load balancing in cloud computing," IOSR Journal of Computer Engineering (IOSR-JCE), vol. 16, no. 4, pp. 23–28, 2014.

[22] J. Kaur and S. Kinger, "A survey on load balancing techniques in cloud computing," International Journal of Science and Research (IJSR), vol. 3, no. 6, pp. 2662–2665, 2014.

[23] M. Randles, D. Lamb, and a. Taleb-Bendiab, "A Comparative study into distributed load balancing algorithms for cloud computing," IEEE 24th Int. Conf. Adv. Inf. Netw. Appl. Work., pp. 551–556, 2010.

[24] J. Uma, V. Ramasamy, A. Kaleeswaran "Load Balancing Algorithms in Cloud Computing Environment - A Methodical Comparison " International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), Volume 3, Issue 2, February 2014