Volume: 08, Issue: 05 Pages: 67-71 (2017) Special Issue

A Survey on Load Balancing in Cloud Computing using Various Algorithms

G.Angayarkanni
Department of Computer Science, TBAK College for Women, Kilakarai
Email: g.angayarkanni@gmail.com

ABSTRACT

Cloud Computing helps to share knowledge and supply several resources to users. Users pay just for those resources the maximum amount they used. Cloud Computing stores the information and distributed resources within the open surroundings. The number of knowledge storage will increase in open surroundings. So, Load balancing is one among the most challenges in cloud computing that is needed to distribute the workload equally across all the nodes. Several algorithms have developed for allocating client's requests to accessible remote nodes. Efficient load balancing ensures dynamic resource utilization of resources to customers on demand basis and increased the general performance of the cloud. This paper may be a transient discussion on the present load balancing techniques in cloud computing and any compares them supported varied parameters like processing time and interval etc.

Keywords - Cloud Computing, Load Balancing, Dynamic, Resource Utilization.

1. Introduction

Cloud Computing [1] is an attracting technology within the field of engineering science. The cloud is dynamical our life by providing users with new varieties of services. Users get service from a cloud while not taking note to the small print. Agency gave a definition of cloud computing as a model for facultative omnipresent, convenient, on-demand network access to a shared pool of configurable computing resources (eg., networks, servers, storage, applications, and services) that may be apace provisioned and discharged with small management effort or service supplier interaction and a lot of individuals listen towards cloud computing.

Cloud computing is economical and ascendable however maintaining the soundness of process such a lot of jobs or requests within the cloud computing atmosphere could be a terribly complicated downside. So the job arrival pattern isn't foreseeable and therefore the capacities of every node within summary of those algorithms and discuss their properties.

The load balancing model is aimed toward the general public cloud that has various nodes with distributed computing resources in many alternative geographic locations. This model divides the general public cloud into many cloud partitions. Once the atmosphere is incredibly giant and sophisticated, these divisions modify the load balancing.

The cloud incorporates a main controller that chooses the appropriate partitions for inward jobs whereas the balancer for every cloud partitions chooses the simplest load balancing strategy main controller and balancer helps to balance the load and to boost the potency. During this paper we tend to gift a survey of this load balancing algorithm developed specifically to suit the cloud computing environments.

2. Load Balancing

Load balancing [1] are often outline as a way for distributing work on the multiple laptops or a computer cluster through network links to attain best resource utilization that maximizes turnout and minimizes overall time interval. It minimizes the overall waiting time of the resources moreover as avoids an excessive amount of overload on the resources. In this technique traffic is split among servers, in order that knowledge are often sent and received while not most delay. One of the crucial issue in cloud computing is to divide the work dynamically. Work of a machine suggests that the overall time interval it needs to execute all the tasks allotted to the machine. Load balancing is that the method of up the performance of the system by shifting of work among the processors. The advantages of distributing the work include higher resource utilization quantitative relation that more results in enhancing the performance thereby achieving most consumer satisfaction.

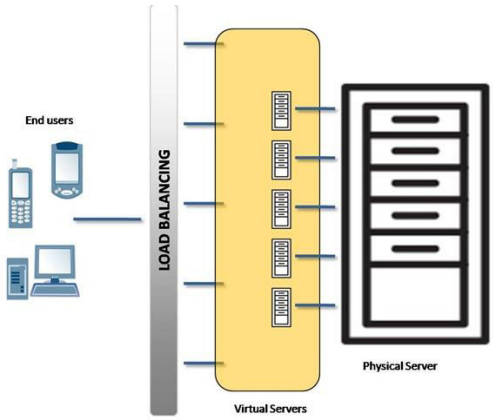

Fig: Load Balancing in cloud computing

3. Analysis of Various Existing Load Balancing Algorithm

3.1 Static Load Balancing

Round Robin Algorithm [2]: The round-robin load balancing algorithm uses the round-robin strategy for allocating jobs. It selects the primary node arbitrarily and then allocates jobs to all or any alternative nodes in a round robin fashion. Processors area unit allotted to every method in an exceedingly circular order with none variety of priority and then there's no starvation.

Weighted Round Robin Algorithm: Weighted round-robin was developed to enhance the crucial issues with round robin algorithm. In weighted round robin algorithm, every server is allotted a weight and in line with the values of the weights, jobs area unit distributed. Processors with bigger capacities are assigned a larger value. Therefore the best weighted server can receive a lot of tasks. In a very state of affairs wherever all weights become equal, servers can receive balanced traffic.

Min-Min Load Balancing Algorithm [3]: It starts with a group of all unassigned tasks. During this minimum completion time for all tasks is found. Then after that among these minimum times the minimum worth is chosen. Then task with minimum time schedule on machine. After that the execution time for all different tasks is updated on that machine although identical procedure is followed till all the tasks area unit allotted on the resources.

Max-Min Load Balancing Algorithm [3]: Max-min load balancing algorithm is analogous to the min-min algorithm except the following: when looking for the minimum execution times, the maximum value is chosen which is the maximum time amongst all tasks on the resources. Then in step with the most time, the task is scheduled on the corresponding machine. The execution time for all different tasks is updated on it machine and therefore the allotted task is aloof from the list of tasks that area unit to be allotted to the machines.

Opportunistic Load Balancing Algorithm [4]: In a cloud environment, every host as a process node performs a task or a subtask. The OLB algorithm intends to stay every node busy despite the present employment of every node. OLB assigns tasks to obtainable nodes in random order.

Shortest Job Scheduling Algorithm [5]: In this algorithm shortest practicable job is chosen first. The approach follows to perform the entire execution of short jobs to utilize the resources in completion of serious jobs. Shortest job had a benefit that the waiting time for the processes is a smaller amount that makes it a robust approach.

Two-phase (OLB + LBMM) load balancing algorithm [6]: This scheduling algorithm is to realize higher executing efficiency of the system. Working procedure of OLB algorithm is to place each and every single node in operating condition so that the goal of cloud computing can be achieved. On the opposite hand LBMM algorithm is employed for minimizing the execution time of the tasks on node that scale back of overall completion time. Combining these two algorithms facilities achieving proper utilization of all resources and enhances the work potency within the network of multiple processor.

Central load balancing policy for virtual machines (CLBVM) [7]**:** CLBVM proposed a policy that balances the load equally in a distributed virtual machines/cloud computing atmosphere. This policy improves the performance of the system however doesn't take into account the systems that area unit fault-tolerant.

Active Monitoring Load Balancing (Optimal) Algorithm: In equally spread current execution the random arrival of load during a cloud setting will reason behind some the server to be heavily loaded whereas alternative server is idle or lightly loaded. Equally load distributing improves performance by transferring load from heavily loaded server to gently loaded servers. Here the jobs are submitted by the clients to the automatic data processing system. Because the submitted jobs arrive to the cloud they're queued within the stack. The cloud manager estimates the duty size and checks for the supply of the virtual machine. Once the duty size match, the job scheduler instantly allocates the identified resources to the duty in queue.

Equally Spread Current Execution Algorithm: The random arrival of load in such associate surroundings will cause some server to be heavily loaded while other server is idle or only gently loaded. Equally load distributing improves performance by transferring load from heavily loaded server. Powerful scheduling and resource allocation may be a vital characteristic of cloud computing based on which the performance of the system is calculable. The considered have a bearing on cost optimization, which may be obtained by improved latent period and interval.

3.2 Dynamic Load Balancing Algorithm

Power Aware Load Balancing [8]: In this process first of all utilization share of every computing node is calculable for the working module, that decides the number of operating computing nodes whereas different nodes are completely stop or not in working condition. This algorithm has 3 sections in operating module: balance section, upscale section and downscale section. Balance section is responsible for finding initialization process wherever virtual machine goes to begin. The second section power-up the extra computing nodes and third downscale section shut-downs the idle work out node in the process participant.

Fuzzy Active Monitoring Load Balancing: Srinivas Seth et al. [9] proposed a load balancing algorithm based on fuzzy logic and uses two parameters processor speed and weight on virtual machine. In [10], the authors have introduced a new fuzzy logic based dynamic load balancing algorithm with additional parameters like disk space usage, , memory usage, bandwidth usage, virtual machine status and named it as fuzzy Active Monitoring Load Balancer.

Throttled Load Balancing: In the paper [11] author defines algorithm in which the client first requests the load balancer to detect a suitable virtual machine to perform the essential operation for the incoming process. In Cloud computing, there may be several instances of virtual machine. These virtual machines can be combined based on the type of requests they can handle. So as per the incoming requests it works accordingly. Whenever a client sends a request, the load balancer will first view for that group and if it is ready to accept and handle the request it is going to select request to it.

Honeybee Foraging Behavior: It is a nature excited Algorithm for self-organization. Honeybee attains global load balancing through local server actions. The performance of the system is enlarging with increased system diversity. The main problem is that throughput is not increased with an increase in system size. When the distant population of service types is required then this algorithm is best suited.

Active Clustering - In this algorithm similar type nodes of the system are grouped together and they work together in groups. This works like as self-aggregation load balancing technique where a network is rewired to balance the load of the system. Systems optimize using identical job assignments by connecting similar services.

Biased Random Sampling: This algorithm is based on the development of the virtual graph having combination between the all nodes of the system where each node of the graph is like to the node computer of the cloud system. Edges between nodes are two types as Outgoing edge and incoming edge that is used to examine the weight of particular system and also allotment the resources of the node. It is very nice technique to balance the load.

Compare and Balance - This algorithm is uses to reach equilibrium condition and handle unbalanced systems load. On the basis of probability, current host arbitrarily select a host and analyze their load. If load of current host is greater than the selected host, it shift extra load to that particular node. Then each host of the system performs the similar procedure. This load balancing algorithm is also designed and implemented to decrease virtual machines migration time.

Shortest Response Time First: In this every process is assigned a priority which is grant to run. In this equal priority processes are scheduled in FCFS order. In this algorithm the priority is the inverse of the next CPU burst. It means, if greater the CPU burst then lower the priority. The SJF policy choose the job with the shortest (expected) processing time first. In this algorithm shorter jobs are executed before long jobs.

Ant Colony Optimization: This algorithm is a multiagent approach to difficult combinatorial optimization problems. Example of this approach is quadratic assignment problem and the travelling salesman problem. These algorithms were influenced by the observation of real ant colonies. Ant's behavior is directed larger to the survival of the colonies. They not think for individual.

Genetic Algorithm Based Load Balancing: It is proposed by Brototi Mandal Kousik and Dasgupta [12]. This algorithm flourish to balance the load of the cloud infrastructure while trying to decrease make span of a job. Genetic based approach follow a few rules and randomization according to the network load effectively.

Generalized Priority Algorithm: In this algorithm [14] the tasks are preferred according to the size of the tasks such that the task with greater size gets the highest priority in the system and execute first at its finest. Also the virtual servers are preferred according to their million instructions per second (MIPS) value in the virtual server distribution system, such that the Server with the greatest MIPS value gets highest priority. Hence the load balancing is done accordingly and it gains the maximum utilization of the resources according to the data size in progress.

Load Balancing Based On A Lock-Free Multiprocessing Solution : X. Liu et al. [13] proposed a technique that skips the use of shared memory whereas other multiprocessing load balancing solutions used shared memory concept and locks to sustain a user session. So that this technique is called a lock-free multiprocessing solution for load balancing. This solution helps in developing the overall performance of load balancer in a multi-core environment by running many load-balancing processes in one load balancer.

Join-Idle-Queue: Y. Lua et al. [15] proposed an algorithm for systems and web services called as Join-Idle-Queue load balancing algorithm. It facilitates large scale load equalizer with distributed dispatchers. In each dispatch firstly load equalizer algorithm idles the processors for the availability and then does allocation of the task to processors in such a way that decreases the queue length at each server. This algorithm remove the load balancing work from critical path of request processing which helps in active reduction of the system load.

Stochastic Hill Climbing Technique: Kousik Dasgupta et al. proposed [16] a novel load balancing strategy by using Stochastic Hill Climbing algorithm. The hill climbing chooses arbitrarily form the uphill moves with effective probability. Author of a local optimization

approach Stochastic Hill climbing utilize the resources and the algorithm is used for allotment of incoming jobs to the servers or virtual machines (VMs).

Decentralized Content Aware Load Balancing: H.Mehta et al. [17] proposed a different content aware load balancing policy known as Workload and Client Aware Policy. This policy explains the unique and special property of the requests as well as computing nodes by a exclusive and special property. USP helps the scheduler in deciding the good and fit resources to implement the process. This technique has lower overhead and implemented in decentralized manner. This technique increases the searching performance by using the content information. It also develop the utilization of resources by reducing the idle time of the computing nodes.

Server-Load Balancing For Internet Distributed Services: A. M. Nakai et al. [18] proposed a distributed server based technique for web servers. It promotes the reduction in service response time by using a protocol that forced the redirection of requests to the closest remote servers without overloading them. A middleware is used in this technique to implement this protocol. To tolerate overload, web server uses Heuristic.

Load Balancing Based On A Lock-Free Multiprocessing Solution : X. Liu et al. [19] presented a technique that neglects the use of shared memory whereas other multiprocessing load balancing solutions used shared memory concept and locks to keep a user session. That's why this technique is called a lock-free multiprocessing solution for load balancing. This solution benefits in improving the overall performance of load balancer in a multi-core environment by running numerous load-balancing processes in one load balancer.

Load Balancing Scheduling for Virtual Machine Resources: Hao Liu et al. [20] presented a scheduling strategy that uses current status of the server and previous logs. By using a genetic algorithm approach this strategy helps in reduction of dynamic migration in the system. It helps in resolving the problem of load-imbalance and high cost of migration thus achieving better resource utilization.

Load Balancing strategy for Virtual Storage (LBVS): H. Liu et al. [21] presented a load balancing virtual storage strategy that provides a storage as a service model and large scale net data storage model based on Cloud storage. Storage virtualization is attained using an architecture that is three-layered and load balancing is attained using two load balancing modules. It improves the efficiency of concurrent access by using replica balancing which decreases the response time and enhances the capacity for backup.

Load Balancing based on a Task Scheduling Algorithm: Y. Fang et al. [22] presented a two-level task scheduling Mechanism. It is used for contest the dynamic requirements of users and obtaining a high utilization of resources. The mechanism is based on load balancing. In this mechanism firstly tasks are first mapped to virtual machines and then virtual machine to host resources. This mechanism improves the resource utilization as well as task response time.

4. Conclusion

Cloud computing provides everything to the user as a service over network. The major problem of cloud computing is Load Balancing. Overloading of a system may point to worst performance which can make the technology unsuccessful, for the powerful utilization of resources; the powerful load balancing algorithm is required. In this paper, we have surveyed various load balancing algorithms in the Cloud computing environment. We have discussed the already proposed algorithms by various researchers.

References

[1] Rajwinder Kaur and Pawan Luthra, "Load Balancing in Cloud Computing", Association of Computer Electronics and Electrical Engineers, ITC. , 2014

[2] Nusrat Pasha, Dr. Amit Agarwal[1] and Dr. Ravi Rastogi[2], "Round Robin Approach for VM Load Balancing Algorithm in Cloud Computing Environment", ijarcsse, volume 4, issue 5, May 2014, pg 34- 39.

[3] Amandeep, V. Yadav, F. Mohammad. "Different Strategies for Load Balancing in Cloud Computing Environment: a critical Study." IJSRE, Vol. 3, pp. 85-90, April 2014.

[4] D. Saranya, L. S. Maheswari. "Load Balancing Algorithms in Cloud Computing: A Review." International Journal of Advanced Research in Computer Science and software Engineering, Vol. 5, No. 7,pp. 1107-1111, July 2015.

[5] Poonam Devi,Mr. Trilok "GabImplementation of cloud computing by using short job scheduling" IJARCSSE.

[6] S. Wang, K. Van, W. Liao, and S. Wang, "Towards a Load Balancing in a Three-level Cloud Computing Network", Proceedings of the 3rd IEEE International Conference on Computer Science and Information Technology (ICC SIT), Chengdu, China, pp.108-113, September 2010.

[7] A. Bhadani and S. Chaudhary, "performance evaluation of web servers using central load balancing policy over virtual machine on cloud", proceedings of third Annual ACM.

[8] J. M. Galloway, K. L. Smith, and S. S. Vrbsky, "Power aware load balancing for cloud computing," in Proceedings of the World Congress on Engineering and Computer Science, vol. 1, pp.19–21, 2011.

[9] S. Sethi, A. Sahu, and S. K. Jena, “Efficient load balancing in cloud computing using fuzzy logic,” IOSR Journal of Engineering, vol. 2, no. 7, pp.65–71, 2012.

[10] Z. Nine, M. SQ, M. Azad, A. Kalam, S. Abdullah and R. M. Rahman, “fuzzy logic based dynamic load balancing in virtualized data centers” In fuzzy system (FUZZ), IEEE International conference on, pp. 1-7, 2013.

[11] Ms.Nitika, Ms.Shaveta, Mr. Gaurav Raj; “Comparative Analysis of Load Balancing Algorithms in Cloud Computing”, IJARCET Volume 1, Issue 3, May 2012.

[12] Kousik Dasgupta, Brototi Mandal, Paramartha Dutta, “Load Balancing in Cloud Computing using Stochastic Hill Climbing-A Soft Computing Approach”,Elsevier (C3IT) 2012.

[13] Hu, 1. Gu, G. Sun, and T. Zhao, "A Scheduling Strategy on Load Balancing of Virtual Machine Resources in Cloud Computing Environment", Third International symposium on (PAAP), pp.89-96,2010.

[15] Kousik Dasgupta, Brototi Mandal, Paramartha Dutta,Jyotsna Kumar Mandal,Santanu Dam , “A Genetic Algorithm (GA) based Load Balancing Strategy for Cloud Computing”, Elsevier (CIMTA) 2013.

[16] H. Mehta, P. Kanungo, and M. Chandwani,"Decentralized content aware load balancing algorithm for distributed computing environments", Proceedings of the International Conference Workshop on Emerging Trends in Technology (lCWET),pp.370-375, February 2011.

[17] A. M. Nakai, E. Madeira, and L. E. Buzato,"Load Balancing for Internet Distributed Services Using Limited Redirection Rates", 5th IEEE LADC ,pp.156-165 2011.

[18] Xi. Liu, Lei. Pan, Chong-Jun. Wang, and JunYuan. Xie, "A Lock-Free Solution for Load Balancing in Multi-Core Environment", 3rd IEEE Int. Workshop on lSA, pp.1-4 2011.

[19] Hu, 1. Gu, G. Sun, and T. Zhao, "A Scheduling Strategy on Load Balancing of Virtual Machine Resources in Cloud Computing Environment", Third International symposium on parallel architecture, algorithms and programming(PAAP),pp.89-96,2010.

[20] Hao Liu, Shijun Liu, Xiangxu Meng, Chengwei Yang, Yong Zhang, LBVS:A Load Balancing Strategy for Virtual Storage, IEEE International Conference on Service Sciences, 2010.

[21] Y. Fang, F. Wang, and J. Ge, "A Task Scheduling Algorithm Based on Load Balancing in Cloud Computing", Web Information Systems and Mining, Lecture Notes in Computer Science, Vol. 6318,pp.271-277,2010.

[22] Z. Zhang, and X. Zhang, "A Load Balancing Mechanism Based on Ant Colony and Complex Network Theory in Open Cloud Computing Federation", Proceedings of 2nd International Conference on Industrial Mechatronics and Automation (ICIMA), Wuhan, China,pp.240- 243, May 2010.