ISSN:1991-8178

Australian Journal of Basic and Applied Sciences

Journal home page: www.ajbasweb.com

Analysis of Load balancing Algorithms in Cloud Computing

Abinaya Pasupathi,R. Saranya, G. Prabhakaran, T. Rajasekaran

Department of Computer Science and Engineering Saveetha School of Engineering Saveetha University Chennai, India

ARTICLE INFO

Article history:
Received 12 March 2015
Accepted 28 April 2015
Available online 1 June 2015

Keywords:
Green Computing, Virtualization, Migration, Load Balancer, Load balancing.

ABSTRACT

Cloud Computing could be a pool of resources that may be shared among the users. At present, Cloud Computing is a rising technology since it provides services at the user level. Users submit larger tasks to the cloud that performs the task by using the servers and also theresult's given back to the corresponding user. There aremanyproblems or challenges in the Cloud Computing atmospherelikeconvenience, Security, and Resource Allocation etc., this paper concentrates on convenience of nodes within the cloud. Balancing Load underneath the nodes can increase convenience of nodes in Cloud. To enhance the performance of the whole cloud atmosphere, efficient Load equalization techniques arerequired. Load Balancing (LB) algorithms distribute the load equally across allthe nodes in cloud. Load balancing in Cloud Computing can increase the liableness and user satisfaction. This survey paper comparesnumerous Load balancing algorithms in cloud.

© 2015 AENSI Publisher All rights reserved.

To Cite This Article: Abinaya Pasupathi,R. Saranya, G. Prabhakaran, T. Rajasekaran, Analysis of Load balancing Algorithms in Cloud Computing. ***Aust. J. Basic & Appl. Sci.,*** *9(11): 605-610, 2015*

INTRODUCTION

All the massive business and tiny business corporationssquare measure moving to cloud atmosphere thanks to its measurability. The rolesinbound to the Cloud atmospheresquare measuredead by the massiveknowledge centers that have thousands of blade servers. Cloud Computing could be a Service destineddesign (SOA). It provides differing kinds of services to the users. Users will get the services with no have to be compelled torecognize their infrastructure. That is, users don'trecognizewherever the service is originated andits infrastructure. Users have to be compelled to pay just for what they used from cloud within thesort of services. This can be the simplicity of Cloud. The most effective example of Cloud is that theelectrical phenomenonprovides that we square measuremistreatment in our day-after-day life. The facility is created by massive wind mills or away. Then the created power is remodeledto varied areas by means that of wires and transformers. Users consume these power and pay just for what they used. The facilitycustomersdon'trecognizethe small print of whereverthe facility is createdand the wayit'sremodeled to their homes. There square measure fourdiffering kinds of cloud atmosphere. They're Public Cloud (Free of price, anyone will access), non-public Cloud (Pay for what you used, just for single organization people), Hybrid Cloud (Combined each public and non-public Clouds), community cloud (For Communication purpose) India. Users will access the cloud resources within thesort ofamenities. There square measurethree basic services provided by the Cloud atmosphere. They're Platform as a Service (PaaS), code as a Service (SaaS) and Infrastructure as a Service (IaaS)

Virtualization:

Cloud Computing is predicated on the construct of virtualization technology. Virtualization means that "something that's not real howeveroffers all the facilities as a true one". It's the codeexecution on the vacant hardware in order that the resources underneath the hardware will beutilizeda lot of effectively. Cloud Computing uses the virtualization technique to create uses the cloud resources with efficiency (Ratan Mishra 2012). There are two styles of virtualization will beemployed in cloud atmosphere.

Full Virtualization:
Para Virtualization:
Full virtualization:

In Full Virtualization (Ratan Mishra 2012), the installation of 1laptopis finished on the oppositelaptop. It'llend invirtual machines that have all the facilities and software that square measuregiftwithin the actual machine.

Corresponding Author: Abinaya Pasupathi, Assistant Professor Department of Computer Science and Engineering Saveetha School of Engineering Saveetha University Chennai, India.
E-mail: abipasu5530@gmail.com

Para virtualization:

In Para Virtualization (Ratan Mishra 2012), the hardware permits multiple operational systems to be run on one machine. For this, Virtual Box tool is employed. Here all the services don't seem to beabsolutelyon the market, instead of the services square measure provided in a very partial manner.

Live migration:

Live Migration (Liang, 2009) could be a virtualization technique which will increase the resource utilization overthe complete and Para Virtualization strategies. During thismethodology, the Virtual machine is migrated from one physical machine to alternativewhile not halting the presently running programs. Inancient techniques, whereasplayacting migration, all the programs have to be compelled toispack up then solely virtual machine will be migrated. However with Live Migration technique, all the presently running programs can continue their operation. Thereby increasing the resource utilization and turnout. Inexperienced Computing Green Computing (Srimathi, 2012) is predicated on the itemswhich will not damage any natural atmosphere and possesses less damage to the atmosphere. Since all the littleand enormous business vendors were moving to cloud atmosphere, the usage of knowledge center is raised. The facility consumed by these knowledge centers can even be raised. So asto realizeinexperienced Computing, the facility consumption is to be reduced. Power Consumption is directly associated with Carbon Emission. If the Carbon Emission is reduced then power consumption is additionally be reduced.

Load balancing:

Load balancing (Gaochao, 2013) could be a technique to balance the load across cloud atmosphere. It'sthe method of shifting load from heavily loaded nodes to low loaded nodes. As a result, no node ought to be heavily loaded. Thereby it'llsurgethe supply of nodes.

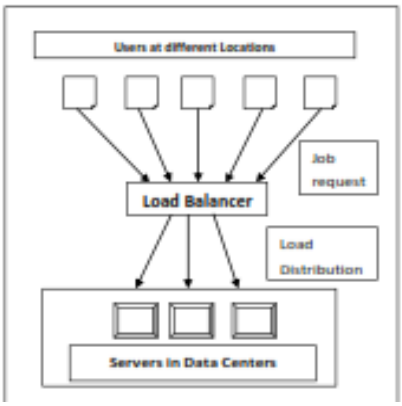

Fig. 1: shows the foremost works of Load balancing technique. The Load Balancer is also any codethat receives jobs from totally different users in several locations. The received massesare distributed equally across all the servers in information Center.

Need for load balancing:

If all the rolesare arrived to the onlyone node, then its queue size is increased and it becomes overladen. There'sa desire to balance the load across numerous nodes, in order thatevery node is in running state however not in overladen state. The goals are as follows (Zenon Chaczko, 2011):

- To increase the supply
- To increase the user satisfaction
- To improve the resource deploymentmagnitude relation
- To minimalize the waiting time of job in queue furthermoreoncut back job executiontime
- To improve the performance of Cloud atmosphere

Basic types of load balancing algorithms:

Depending on the instigator of the algorithmic rule, Load balancing algorithms may beclassified into 3sorts (Ratan Mishra 2012): Sender Initiated Sender identifies that the nodes areoverpoweredinorder that the sender initiates the execution of Load balancingalgorithmic rule.Receiver initiated the requirement of Load balancingscenariomay beknown by the receiver/serverin cloud which server initiates the execution of Load balancingalgorithmic rule.Symmetric-it is the mixture of each the sender initiated and receiver initiated kinds. It takes benefits ofeach sort.Based on this state of the system, load balancing algorithms may be divided into 2 types:

Static Schemes-The current standing of the node isn't taken under consideration (Venubabu Kunamneni, 2012). All the nodes and theirproperties are predefined. Since itdoesn't use current system standingdata, it's less advanced and it'sstraightforward to implement.

Dynamic Schemes-This type of algorithmic rulerelies on this system data (Venubabu Kunamneni, 2012) and the algorithmic rule works in keeping with the changes within the state of nodes. Dynamic schemes arepricey one and areterriblyadvanced to implement however it balances the load in effective manner.

Status Table-Status table (Gaochao, 2013) could be a arrangement to take care of this standing of all the nodes within thecloud atmosphere. This datamay beemployed bya number of the dynamic theme algorithms toallot jobs to the nodes that aren't heavily loaded.

Load balancing algorithms:

Several Load balancing algorithms were projected. a number of those algorithms are mentioned here.

Dynamic spherical robin algorithmic rule:

It is Associate in nursing extension to the spherical Robin algorithmic rule (Karanpreet Kaur, 2013). Every physical machine has range of virtual machines. Thealgorithmic ruleschiefly works on reducesthe facility consumptionof physical machine that, it uses 2 rules:

1. The virtual machine in one physical machine has finished its execution however all different virtual machines are still running then if any new virtual machine is arrived suggests that the corresponding physical machine doesn'tsettle for the new virtual machine. Such physical machines areknown as "retiring" state physical machines (i.e.), once all the opposite virtual machines has finished their execution then we willclean up that physical machine.
2. The second rule says that if a physical machine is in retiring state for an extended time then rather than waiting, all the running virtual machines are migrated to different physical machine.This algorithmic rule works to save lots ofa lot of power by clean up the physical machines and thusthe price for power consumption is little. Even though it doesnotproportionfor a giantinformation centers.

Hybrid algorithmic rule:

It is the mixture of each Dynamic spherical Robin and initial - match algorithms (Gaochao, 2013). Itsaccustomedcut backthe facility consumption by physical machines. the primary - matchalgorithmic rule is applied throughout rush hours to extend resource utilization of physical machines and Dynamic spherical Robin algorithmic rule is applied throughout non-rush hours to consolidate virtual machines and clean up the physical machines. This algorithmic ruleis employed for up Resource allocation and Power Consumption and this algorithmic ruledoesn'tproportionfor giantinformation centers.

Equally spread current execution (esce) algorithmic rule:

Cloud Manager estimates the dimensionsof every job within the queue and looking out for thesupply of resources for that job (Hemant, 2013). If all the resources areon the market for that job then straightawaythe taskhardware allocates that job to the resource. This algorithmic rule is for uplatency and time interval of employment. Howeverit's not fault tolerant and it'sthe matter of single purpose of failure.

Central load balancing policy for virtual machine:

This algorithmic rule balances load equally across the distributed systems and cloud computing environments (Shreyas Mulay, 2013) and this policy recoversthe performance of the system howeverdoesn'tthink about the systems that are fault-tolerant.

Enhancing equally distributed load balancingalgorithmic rule:

This algorithmic rule (Shreyas Mulay, 2013) works to distribute the load equally across all the nodes in cloud atmosphere. It handles the requests with priorities. It's a dispersedalgorithmic rule by that the load may be distributed not solelyduring a balanced manner howeveradditionally it allocates the load consistently by checking the counter variable of everyinformation center. When checking, it transfers the load consequently (i.e.) minimum price of the counter variable are chosen and also the request is handledsimply and takes less time and providesmostoutturn. For every arrival of job, the counter variable is augmented by one and for dispatching a job; the counter variable is weakened by one. The algorithmic rule is dead by one central server node. During thismethodology, the central server node gets overhead by receiving a lot of jobs. The algorithmic rule allocates job supported the counter variable and it doesn'tthink aboutthe burden of the task.

Decentralized content aware load balancingalgorithmic rule:

It is supportedwork and consumer Aware Policy (WCAP) (Nidhi Jain Kansal, 2012). It uses idiosyncraticsingular noveland Special Property (USP) to specify the unique and special property of the requests furthermore as computing nodes. USP helps the hardwareto make a decisionthe most effectiveappropriate node for process the requests and this strategy is obligatoryduring alocalized manner with low overhead and thus by exploitation the content data to slim down the search, this system improves the looking performance of the system. It additionally helps in reducing the idle time ofthe computing nodes thusup their utilization.

Join-idle-queue:

This algorithmic rule provides giant scale load balancing with distributed dispatchers by,initial load balancing the idle processors across dispatchers for the supply of idle processorsat every dispatcher and so, assignment jobs to the processors to cut back average queue length ateveryprocessor (Nidhi Jain Kansal, 2012) and thus by removing the load balancing work from the vital path of requestprocess, it effectively decreases the system load, acquires no communication slide at job arrivalsand doesn't increase actual latency.

Honeybee foraging behavioralgorithm:

It is a localized honeybee-based load balancing technique that's a nature-inspiredalgorithmic rule for organization [10]. It achieves international load balancing through native server actions. Performance of the system is increased with augmented system diversity howeveroutturnisn'taugmented with a rise in system size. It's best suited to the conditions

whereverthe varied population of service sorts is needed.

Min-min algorithmic rule:

Each job has the execution time and completion time (Karanpreet Kaur, 2013). The cloud manager identifies each theexecution time & completion time of everyspecial job in queue. The taskthat has minimum completion time is knownand so assigns the task to the processor that has the potentialto completethe taskat intervals its fixed completion time. However larger task should be waited forlong amountof your timewithin the queue.

Max-min algorithmic rule:

It works because the Min-Min algorithmic rule. However it providesa lot of priority to the larger tasks (Mohana Priya, 2013). The roles that have giant execution time or giant completion time aredeadinitial. The matter is that smaller jobs got to be waitingdurable.

Rasa algorithmic rule:

It is the mixture of Min-Min and Max-Min Algorithms (Mohana Priya, 2013). The algorithmic rule builds a matrix C wherever C represents the completion time of the task Ti on the resource R_j. If the quantityof obtainable resources is odd and the Min-Min algorithmic rule is useful to assign the primary task and the other side Max-Min algorithmic rule is applied. The remaining tasks areassigned to their acceptable resources by one in every ofthe 2waysinstead. Variousexchanges of Min-Min and Max-Min waysends up in consecutive execution of a tiny low and an oversized task on totally different resources and herewith, the waiting time of the little tasks in Max-Min algorithmic ruleand also the waiting timeof the big tasks in Min-Min algorithmic ruleareunnoticed.

Improved max-min algorithmic rule:

It is Associate in nursing extension to the Max-Min algorithmic rule (Elzeki, 2012). The MaxMinalgorithmic rule selects the task with the most completion time and assigns it to the resource on thatcome through minimum execution time. The fundamentalplan of Associate in nursing improved version of Max-Min algorithmic rule assign task with most execution time to resource produces minimum complete time instead of original Max-Min assign task with most completion time to resource with minimum execution time. It uses the benefits of Max-Min and additionally covers its disadvantages.

Phase load balancingalgorithmic rule:

It is the mixture of OLB (Opportunistic Load Balancing) and LBMM (Load Balance Min- Min) programming algorithms to utilize higherexecutionpotency and maintain the load balancing of the system (Karanpreet Kaur, 2013). OLB programmingalgorithmic rule keeps each node in operating state to attain the goal of load balance and LBMM programmingalgorithmic ruleis usedto reduce the execution of your timeof every task on the node thereby minimizing the completion time. This algorithmic rule works to reinforcethe employment of resources and enhances the work potency.

Power aware load balancing (palb) algorithmic rule:

This algorithmic rule is enforced in cluster controller (Jeffrey Galloway, 2011). There's a presence of Job hardware whose work is to simulate requests from users for virtual machine instances. The cluster controller maintains the employment state of every active calculate node and makes choices on wherever to instantiate new virtual machines and this algorithmic ruleis chiefly designed to cut backthe facility consumption and this algorithmic ruleis working to power off the physical machines that are in idle staterather than getting into low power state. However it doesn'tpartfor giant cloud information centers.

Comparison of loadbalancing algorithms:

Table 1: Show the comparison of LB algorithms which werediscussed above.

Techniques and Algorithms	Specifications	Pros
ESCEAlgorithm	Estimate the size of joband look for availabilityof resources	Improvedresponse time andprocessing time
DynamicRound Robin Algorithm	1. Uses two rules tosave the power consumption 2. Works forconsolidation of VM	Reduce the powerconsumption
HybridAlgorithm	1. Combination ofDynamic Round Robinand First-Fit Algorithm 2. Applied in non-rushhours and rush hours	1. ImprovedResourceUtilization 2. Reduced PowerConsumption
EnhancedEquallyDistributed LoadBalancingAlgorithm	Based on the countervariable, the job is allocated by CentralServer	1. ComputingResource isdistributed efficiently andfairly 2. Reducesrequest toresponse ratio
Central LoadBalancingpolicy forVM	Balances the loadevenly	Improves overallperformance
DecentralizedContentAware LoadBalancingAlgorithm	1. Uses Unique andSpecial Property(USP) of nodes 2. Uses contentinformation to narrow down the search	1.Improves thesearchingperformancehence increasingoverallPerformance 2. Reduces idletime of nodes
Join-IdleQueueAlgorithm	1. Assigns idleprocessors todispatchers for	1. Reducessystem load

	theavailability of idleprocessors 2. Then assigns jobs toprocessors to reduceaverage queue length	2. Lesscommunicationoverhead
Min-MinAlgorithm	1. Estimates minimumexecution time and minimum Completiontime 2. Jobs havingminimum completiontime is executed first	Smaller tasks areexecuted quickly
HoneybeeForagingBehavior	Achieves global loadbalancing through localserver actions	ImprovedScalability
Max-MinAlgorithm	1. Same as Min-Min 2. Gives more priorityto larger tasks than smaller one	Larger tasks areexecuted quickly and efficiently
ImprovedMax-MinAlgorithm	1. Improved version ofMax-Min Algorithm 2. Assigns task withminimum execution time	Scheduling jobsEffectively
RASAAlgorithm	Combination of bothMin-Min and Max-MinAlgorithms	1. Efficientresourceallocation 2. Minimumexecution time
2-Phase LoadBalancingAlgorithm	1. Uses OLB to keepeach node busy 2. Uses LBMM toachieve minimum execution time of eachjob	1. Efficientutilization ofresources 2. Enhances workEfficiency
PALBAlgorithm	1. Implemented inCluster Controller 2. Use Job Scheduler tosimulate requests fromusers for virtualmachine instances	PhysicalMachines that arein idle state aremove to poweroff state to conserve energy

Proposed system:

Existing Load balancing algorithms have some drawbacks in up overall performance of the cloud atmosphere. Still there'sa retardant of overloading nodes within the Cloud atmosphere.It’sdreadfullytroublesome to manage entire cloud atmosphere. Thus the projectedplan is to divide the whole cloud atmosphere into many partitions supported its geographical locations (Gaochao, 2013). Currently the Load balancingalgorithmic rulemay be applied solely to the partitions andto notthewhole cloud. Fig a pair of shows the cloud atmospherewhen partitioning is completed. The loadbalancingalgorithmic rule is applied to every partition so as to avoid overloading of nodes.For balancing load in cloud partition model (Gaochao, 2013) there aretwo vitalparts needed:Load Balancer and Main Controller. Load Balancer is related toevery partition whose work is to take care of the state data and is to be updated in periodic intervals. Then the task is allotted to the partition if it's in idle or traditional state and whenassignment job to the partition, the balancer should update the standingdataof every node in this partition. Supported this data, the task is allotted to the nodes. Main controller receives all the roles that arrive from the cloud. Whenever the mostcontrollersreceivethe taskit'sto make a decision and that partition to acceptthe task. Every partition has the state datarelated to it. It should be in idle state, traditional state or heavily loaded state and the state of thenode beyond all partition is ready by considering many parameters of that node and theparameter isalso static or dynamic.

• Static parameters embodyrange of Central processing unit, memory size and speed of the processor or central processing unit.

• Dynamic parameters embodycentral processing unit utilization magnitude relation and memory utilization magnitude relation.

Load of the node is calculated by these parameters. The algorithmic rule is to be designed for the nodes that are idle or traditional and it's to update the standingdataof every node sporadically.Load Balancer in every partition maintains the standing table and the standing table comprisesdataregarding the load of all the nodes in this partition. Supportedthe knowledgewithin thestanding table, Load Balancer agrees the standing of the partition. This table is updated by a Load Balancer sporadically. There’san opportunity for inconsistent of the informationwithin thestanding table. For higherstrength, balancer maintains the 2standing table and every of thatare associated set by the "Flag". The Flag is ready by either "Read" or Write" and one table is about by read while another one is ready by write and the knowledge in table thatis ready by "read" is oftenemployed by the algorithmic rule. "Write" denotes that the table data is being updated and it should not be correct. Once the knowledge is updated within the "write" table then the flag isready to "read".

Future work:

The future work is to develop two algorithms, one for the panels of Cloud atmosphere that are inidle state and another one for the partitions that are in traditional states and switch mechanism is required for applying these twovarious algorithms. If the partition is in idle state one easyalgorithmic rule is to be used and later constant partition will became traditional state and variousalgorithmic rule is to be used. The algorithmic rule designed for traditional state partitions ought to be a lot ofeconomicalin order that it avoids the partition turning intooverladen.

Conclusion:

Though there are many problems in cloud atmosphere, it's been wide adopted by several organizations andindustries. Researchers doseveral

works to resolve those problems. ForLoad balancing issue, the answer is to develop appropriate algorithms that balance the load across the divided cloud atmosphere. Tochange algorithms ought to work consequentlybecause the partition standing changes. It reduces the server overhead, increase outturn, surge performance, cut back server power consumption and additionallyallocate the load across nodes.

REFERNECES

Elzeki, O.M., M.Z. Reshad, M.A. Elsoud, 2012. "Improved MaxMinAlgorithm in Cloud Computing", InternationalJournal of Computer Applications.

Gaochao, Xu., Junjie Pang and Fu. Xiaodong, 2013. "A LoadBalancing Model Based on Cloud Partitioning for thePublic Load", IEEE Transactions on Cloud Computing.

Hemant, Dr., S. Mahalle, Prof. Parag R. Kaveri, Dr. Vinay Chavan, 2013. "Load Balancing on Cloud Data Centers",International Journal of Advanced Research in ComputerScience and Software Engineering.

Jeffrey Galloway, M., L. Karl Smith, S. Susan Vrbsky, 2011. "Power Aware Load Balancing for Cloud Computing",World Congress on Engineering and Computer Science.

Karanpreet Kaur, Ashima Narang, KuldeepKaur, 2013. "LoadBalancing Techniques of Cloud Computing",International Journal of Mathematics and Computer Research.

Liang, Liu., Hao Wang, Xue Liu, Xing Jin, WenBo He, Qing Bo Wang, Ying Chen, 2009. "Green Cloud: A NewArchitecture for Green Data Center", ACM Journal.

Mohana Priya, S., B. Subramani, 2013. "A New Approach forLoad Balancing in Cloud Computing", InternationalJournal of Engineering and Computer Science.

Nidhi Jain Kansal, InderveerChana, 2012. "Cloud LoadBalancing Techniques: A Step Towards GreenComputing", International Journal of Computer Science.

Ratan Mishra and AnantJaiswal, 2012. "Ant Colony Optimization: A solution ofLoad Balancing in Cloud", International Journal of Web& Semantic Technology (IJWesT).

Shreyas Mulay, Sanjay Jain, 2013. "Enhanced EquallyDistributed Load Balancing Algorithm for CloudComputing" International Journal of Research inEngineering and Technology.

Srimathi, V., D. Hemalatha, R. Balachander, 2012. "Green CloudEnvironmental Infrastructure" , International Journal ofEngineering And Computer Science.

Venubabu Kunamneni, 2012. "Dynamic Load Balancing forthe cloud", International Journal of Computer Scienceand Electrical Engineering.

Zenon Chaczko, Venkatesh Mahadevan, Shahrzad Aslanzadehand Christopher Mcdermid, 2011. "Availability and Load Balancing in Cloud Computing",International Conference on Computer and Softwaremodeling IPCSI.