International Journal on Recent Researches In Science, Engineering & Technology
(Division of Computer Science and Engineering)
A Journal Established in early 2000 as National journal and upgraded to International journal in 2013 and is in existence for the last 10 years. It is run by Retired Professors from NIT, Trichy. It is an absolutely free (No processing charges, No publishing charges etc) Journal Indexed in JIR, DIIF and SJIF.
Research Paper
Available online at: www.jrrset.com

ISSN (Print) : 2347-6729
ISSN (Online) : 2348-3105
Volume 4, Issue 1, January 2016.
JIR IF : 2.54
DIIF IF : 1.46
SJIF IF : 1.329

WORKLOAD CONSOLIDATION FOR CLOUD DATA CENTERS WITH GUARANTEED QOS USING REQUEST RENEGING

M.Srinivasu[1], R.Umagowri[2], R.Arunkumar[3]
[1]Professor, Department of Computer Science and Engineering, Mahendra Engineering College, Mahendhirapuri, Namakkal District, Mallasamudram, Tamilnadu, India.
[2,3]Assistant Professor, Department of Computer Science and Engineering, Mahendra Engineering College, Mahendhirapuri, Namakkal District, Mallasamudram, Tamilnadu, India.

Abstract
The difficulties of holding fast to stringent Quality of Service prerequisites, productively use assets, and ration vitality utilization are continually being confronted by Cloud Data Center Service Providers. In an offer to proffer answers for these difficulties, various scientists have proposed shifted arrangements. In any case, there still can't seem to be a comprehensive arrangement that handles every one of these difficulties immediately, as these difficulties are intermittently differentiating. Creators along these lines for the most part concentrate on one at that point try to deal with the bargains on the other(s). To propose another plan for stack adjusting that uses different workload classes to ensure end-to-end QoS while rationing vitality with little trade off on either. Trials were finished utilizing CloudSim toolbox and got comes about demonstrate that our plan beats alternate methodologies both as far as vitality preservation and QoS adherence.
Key Words: Load Balancing Method, Quality of Service (QOS), Energy Protection, Distributed Computing

Introduction

The relative decline in cost of Internet get to and the expansion of savvy gadgets has prompted an increment in workloads at Cloud data center server farms. These expanded workloads with changed prerequisites and a not as much as equivalent increment in asset levels have prompted the need to proficiently use Cloud assets keeping in mind the end goal to viably benefit these workloads and in the meantime profit for the Cloud supplier. There is additionally the desperate need to adjust to principles for green figuring by diminishing general vitality utilization and carbon discharge levels.
One way to deal with vitality preservation is server solidification and multi-tenure [1], [2]; which using virtualization and virtual machines [3] look to totals workloads on Physical Machines (PMs) together in an offer to diminish the aggregate number of dynamic PMs. Doing this however could effects on client workloads as the dream of committed PMs which Virtual Machines (VMs) give to clients is in not immaculate and shared assets can be savagely challenged for by these workloads [2]. Cloud suppliers are then confronted with the issue of fighting with vitality protection as opposed to ensuring QoS adherence. In the work, an approach uses numerous workload classes to ensure an end-to-end QoS adherence while in the meantime rationing vitality is proposed.

Related Works

Classification of customer workloads has been done by different makers some of which include: [4] where customer workloads were part into two social events – Gold and Bronze in perspective of

customer required response times. In advance of [5], customer workloads were accumulated into three social affairs – Short, Medium and Long based on the customer demonstrated burst time of every endeavors. In works done by [6], [7] the makers used distinctive customer gave criteria to portrayal of workloads. Regardless of the way that their works focused on workload appointment, they expected to gather these workloads with a particular ultimate objective to choose need of securing. Reference [8] proposed a benefit based course of action of PMs using RAM, CPU and Bandwidth, in which customer workloads were relegated to the PM that offered slightest completing time for such endeavors.

In advance of [9], [10] distinctive SLA parameters, (for instance, thing sort, account sort, request sort, response time etcetera.) were considered by the day's end workloads were organized into three social affairs – Small, Medium and Max or Gold, Silver and Bronze separately. From composing it can be contemplated that gathering of customer workloads is not a silly endeavors, as it is for all intents and purposes hard to consider every essential/worldview in the midst of these courses of action based on the fact that workload portrayal is outside the degree of this work, basically got the best in class approach used.

Reference [11] proposed an energy aware approach to manage errands assignment and load changing in Cloud Data Centers (DC). The focus of this work is on assurance of imperativeness while restricting SLA encroachment. Workloads on affirmation were designated to PMs using a changed best fit dropping estimation called Power-Aware Best Fit Descending (PABFD), which played out a power-advancement test going before workload partition and just administers ensuing to attesting that such dispersion would not make the power usage of the PM more noticeable than a preset edge regard. With respect to stack changing, the approach examines PMs' CPU utilize level against pre-set upper and lower edge regards to recognize an over/under worked. If a PM's CPU use ends up plainly finished as far as possible, VMs are gotten off the PM equivalently, if the CPU utilize is underneath the lower edge, all VMs are gotten off and PM put to rest to screen vitality.

Proposed System

Propose a half breed conspire with highlight sets called Multi-Group Weight Balancer (MG-WB). The proposed approach joins critical upgrades that address the inadequacies of these methodologies while utilizing on their individual qualities. In the main stage, client demands (VMs) are allotted to PMs utilizing our Binary Search Best Fit Algorithm. The proposed calculation is comparative however utilizes the Binary Search Tree (BST) to accelerate the scan for an appropriate PM. It has been demonstrated that BST has a normal, best and most pessimistic scenario running many-sided quality of O(log2n) which for vast passages, is substantially quicker than the normal and most pessimistic scenarios of the straight cluster seek O(n) utilized. Acquaint numerous workload classes with the portion stage. There are three distinct classes of client workloads – Gold, Silver and Bronze and assembled in view of their QoS prerequisites, with Gold being premium and bronze being best exertion.

Like in the work, the use forecast display is utilized as a part of the distribution of VMs to PMs nonetheless; the power development check is expelled. It is normal that the procedure of VM solidification did in the heap adjusting stage would provide food for vitality effectiveness as the higher the quantity of PMs effectively running the higher the aggregate vitality utilization of the whole server farm and the other way around. It is a relationship drawn from the works. Additionally, present a Binary Search Best Fit designation (BSBF), which is utilized as a part of place of the PABFD. PABFD, seeks straightly through every one of the PMs in the server farm for the most appropriate to have a VM. Our defense for this is, given a server farm with N number of PMs, PABFD needs to do N correlation even under the least favorable conditions case before designating a PM to a client workload. On the off chance that N is expansive, this procedure can back off the designation procedure and prompt an expansion in postpone time (SLA infringement). This is the place BSBF has favorable position. Being founded on a twofold pursuit tree, it has a

normal and more regrettable case look unpredictability of log2N consequently ready to discover reasonable PMs Much quicker than direct inquiry based best fit plummeting utilized.
In the load adjusting eliminate the VMs allotment conveyed in the portion stage is enhanced with a perspective of consistently redistributing dispensed workloads among PMs. This would enhance QoS loyalty, and merge VMs onto less PMs to lessen general vitality utilization of the server farm. The stage is part into two sections – usage discovery and VM-Migration. The usage discovery process is the same. In spite of, in picking which VM to move, the class to which it has a place is considered. This suggests all bronze class VMs if exhibit in a PM would be chosen for movement first before any silver class. In like manner all silver classed VMs would be chosen before any gold classed VM is chosen. It would guarantee bring down SLA infringement for the gold class thus unpredictable VM movement. On account of under-used PMs, all VMs are chosen for relocation regardless of the class they have a place with after which the underutilized PMs are placed in rest to ration energy.

Result and Discussion

To check the proficiency of our proposed demonstrate, recreations were completed utilizing CloudSim toolbox and the same trial setup was utilized for correlation reason. The server farm comprised of 800 heterogeneous PMs of two classes and with details and energy consumption models based on benchmarked information from genuine servers.
Table 1 explains the EC (Energy Consumption), AT (Average Time), SLA Violation (Service Level Agreement Violation) for respective input parameters with existing methods. Table 1 displays the average value on all respective evaluation matrix & input parameters with PABFD (Power-Aware Best Fit Descending) and VMCUP existing methods. According to Table1, it noticed that proposed MG-WB algorithm performs well on all evaluation matrix and Input parameters compare than existing methods.

Table.1 Comparison of EC (Energy Consumption), AT (Average Time), SLA Violation (Service Level Agreement Violation)

Algorithm	EC (KWh)	AT (S)	SLA Violation (%)
PABFD	11	1.35	10
VMCUP	10.2	1.0	10
MG-WB	5.2	0.76	7.3

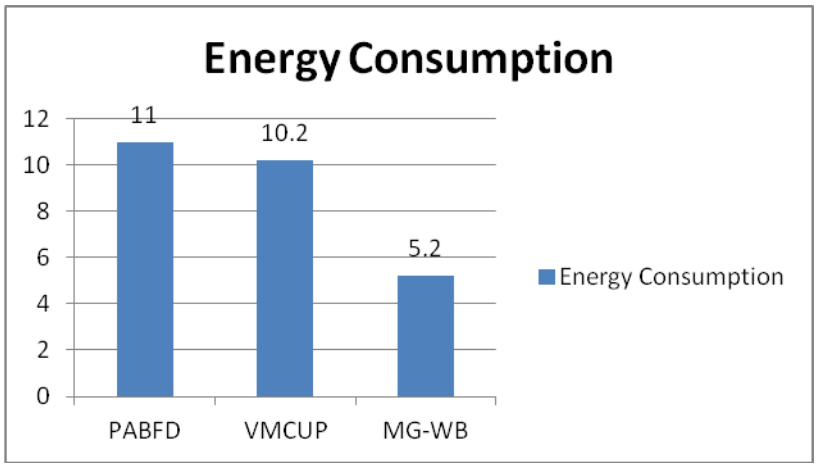

Fig. 1 Comparison of Energy Consumption

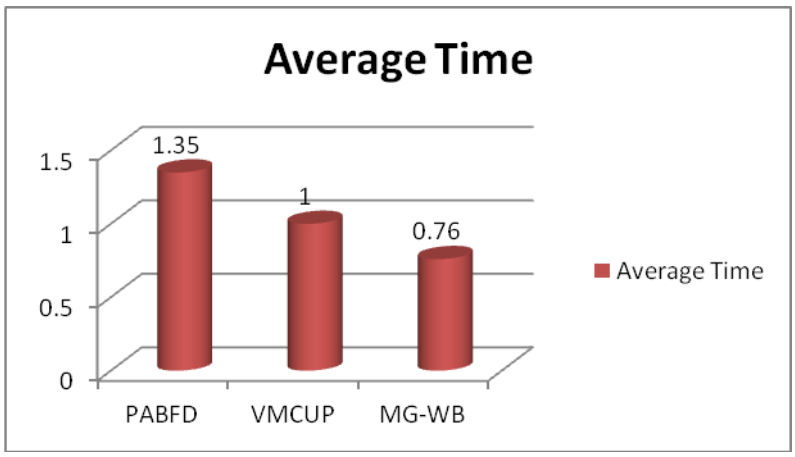

Fig. 2 Comparison of Average Time

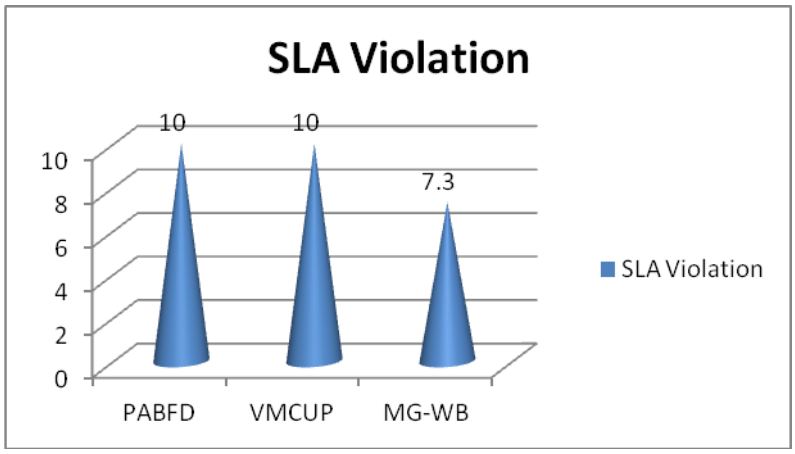

Fig. 3 Comparison of SLA Violation

Fig. 1 to 3 indicates examination of EC (Energy Consumption), AT (Average Time), SLA Violation (Service Level Agreement Violation) for finding efficiency. The proposed MG-WB is estimated with VMCUP and PABFD previous methodologies behalf of EC (Energy Consumption), AT (Average Time), SLA Violation (Service Level Agreement Violation). The proposed methodology is the closest competitor is VMCUP algorithm on overall features. VMCUP is utilized to predict quality of many data sets. However, VMCUP methodology failed to maintain prediction of quality data really slow down. Proposed method reduces the encryption time, decryption time and communication cost in cloud services. Proposed MG-WB reduces 5.0 EC (Energy Consumption), 0.34 AT (Average Time) and 2.7 SLA Violation. Finally, the paper claims the proposed MG-WB algorithm performs best on every evaluation matrix & respective input parameters.

Conclusion

Various research works have been done in asset administration in Cloud processing, however the greater part of them have concentrated on handling a solitary test at any given moment or considering one as the essential test and others as optional. In this work, a way to deal with stack adjusting is suggested that utilizing on the qualities of past works while in the meantime tending to a large portion of their deficiencies is proposed. The proposed approach presents a class-based workload movement combined with a BSBF assignment method. Usage comes about demonstrate that our approach is superior to anything other cutting edge approaches as far as general vitality preservation, SLA adherence and power state exchanging; and marginally less than impressive in the zone of workload delay.

Reference

[1] M. Fiedler, Le-Quoc, C. Cabanilla, "The Top 5 AWS EC2 Performance Problems" Whitepaper. Datadog Inc, 2013

[2] Z. Musgrave, B. Nobel, Y. Xu, B. Nobel, M. Bailey, "Workload-Aware Provisioning in Public Cloud. Internet Computing", vol. 18, no. 4, IEEE Computer Society Press, 2014, pp.15-21.

[3] R. Neugebauer, S. Hand, B. Dragovic, P. Barham, K. Fraser, T. Harris, A. Ho, I. Pratt, and A.Warfield, "Xen and the art of virtualization," Proc. of 19th ACM symposium on Operating systems principles, 2003, pp. 177.

[4] M. Pedram, H. Goudazi, "Multi-dimensional SLA-based Resource Allocation for Multi-tier Cloud Computing Systems" Intl Conf. on Cloud Computing (CLOUD), IEEE Computer Society Press, 2011, pp. 324-331.

[5] E. Ramaraj, A. Karthick, R. Subramanian, "An Efficient Multi Queue Job Scheduling for Cloud Computing", World Congress on Computing and Communication Technologies (WCCCT), IEEE Computer Society Press, 2014, pp. 164-166.

[6] T. Radhika, K. Gouda, M. Akshatha, "Priority Based Resource Allocation Model for Cloud Computing", International Journal of Science, Engineering and Technology Research, vol. 2, no. 1, pp. 215 -219, 2013.

[7] R. Wagh, C. Pawar, "Priority Based Dynamic Resource Allocation in Cloud Computing", Proc. of the Intl Symposium on Cloud and Services Computing, IEEE Cloud Computing, 2012, pp. 1-6.

[8] K. Qian, W. You, Y. Qian, "Hierarchical Queue Based Task Scheduling" Journal of Advances in Computer Networks, 2014, vol. 2, no. 2, pp. 138–141.

[9] S. Garg, L.Wu, R. Buyya, "SLA-based Resource Allocation for Software as a Service Provider (SaaS) in Cloud Computing Environments" 11th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing, 2011, pp. 195-204.

[10] J. Guitart, M. Macias, "Client Classification Policies for SLA Enforcement in Shared Cloud Datacenters", Proc. of 12th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing, 2012, pp. 156-163.

[11] R. Buyya, A. Beloglazov, "Optimal online deterministic algorithms and adaptive heuristics for energy and performance efficient dynamic consolidation of virtual machines in Cloud data centers", Concurrency and Computation: Practice and Experience, pp. 1397–1420.