Indian Journal of Science and Technology, *Vol 8(35), DOI: 10.17485/ijst/2015/v8i35/86650, December 2015*
ISSN (Print) : 0974-6846
ISSN (Online) : 0974-5645

An Enhanced Load Balancing Scheduling Approach on Private Clouds

M. Lavanya[1*], Aarthi Ravi[2], Ajay Aditya[2], Rukmani Samyuktha[2], V. Vaithiyanathan[1] and S. Saravanan[1]

[1]School of Computing, SASTRA University, Thanjavur - 613401, Tamil Nadu, India; m_lavanyass@ict.sastra.edu, vvn@it.sastra.edu, saran@core.sastra.edu

[2]SASTRA University, Thanjavur - 613401, Tamil Nadu, India; aarthi.7395@gmail.com, ajayaditya94@gmail.com, samyutha95@gmail.com

Abstract

This paper aims to bring enhanced load balancing scheduling on private cloud computing. As cloud computing has become a developing and emerging area, the need for a cloud becomes mandatory in working areas. The advantage of technology in their own data centers to evolve a private cloud. Due to the presence of private cloud and to fulfill its unique characteristics and special requirements, it becomes a challenging task to optimally schedule virtual machine instances requests onto nodes to compute when there are multiple tasks to satisfy. In this paper, we discuss about load balancing, which is to schedule task among various end users. The scheduling algorithm used is Round Robin algorithm. This algorithm reduces the response time and improves the processing speed. According to this algorithm the nodes are allocated to the virtual machines along with their corresponding task to a particular node in a cyclic fashion. Each node in the cloud is scheduled equally with the task until all the tasks in the global queue all allocated.

Keywords: Cloud Computing, Load Balancing, Private Cloud, Round Robin Algorithm, Scheduling

1. Introduction

Cloud is a subscription based service used for providing cost based services to its users on demand whenever and wherever needed with storage space and computer resources. For example email service providers such as Gmail, yahoo where the user logins to his/her account in order to access his details rather than having software and storage for his account existent on his computer. User needs internet alone to access the cloud. The components of the cloud include Client machines, Data Centers, Distributed Servers. Client machines are end users who interact with the cloud. Data Centers is the combination or group of servers where application programs reside.

The four types of cloud services are Public-Cloud, Private-Cloud, Community-Cloud and Hybrid-Cloud. The three types of cloud service provided are SaaS (Software as a Service), PaaS (Platform as a Service), and IaaS (Infrastructure as a Service). SaaS provides OS, network, required software for operation. PaaS offers development environment that helps to deploy Java and other language based application. IaaS provides delivery of computer hardware as a service. Windows Azure by Microsoft and Amazon AWS are popular IaaS providers. The applications of Cloud Computing are Search engines (Google, Yahoo and Bing), email sites, social networking sites.

The proposed work is based on few reference papers, among those the paper[1] of Jiandum Li is primary. In this paper they have adapted the load balancing methodology. We use that existing methodology with part-file method and this makes our proposal. In paper[10] dealt with the major problem, the energy consumption of the computer system.

In paper[6] references the key mechanism, the distribution of workloads among data centers dealt with many mechanism and algorithms that enhance

*Author for correspondence

saving power without interfering both workload and performance. Armbrust M, et al[7] emphasizes an efficient description in power management in different datacenter and a decreased Co_2 emission.

In[5] dealt with node group managers and how effectively the virtual machines are allocated to the nodes. In paper[8] show how the workloads are shared in various heterogeneous servers. In Verma A et al[9] make an identification of power management in modern datacenters. In paper[2] establishes the CPU power consumption technique and controlling the carbon emission rate. Round Robin (RR) is the pre-emptive algorithm as the scheduler forces the process based on the time slices and each process are allocated in circular fashion.

In[4] provides the idea on the incoming workloads to servers is multi-structured, hence it contains many sub applications so demand for servers and computer resources is on the hike. So our daily servers need to satisfy energy-use constraints. Thus focusing on conserving power in memory and microprocessor has been established.

In[6] explains on Round Robin as the simplest and easiest way to use algorithm based on time slices. All processes in scheduler is assigned with a time slice which prevents the process starvation. Process starvation does not allow the process to be kept in the queue for long interval, as when the turn occurs for the processes the work begins. Round Robin is applied mostly in network schedulers. Data packet scheduling most probably occurs in computer networks. Time slices in RR really made us easy to move on with all the tasks and in the implementation of Part file methodology.

In[3] explains that the private cloud is widely used in daily scenarios in large organizations to handle their large volumes of data in an efficient manner. But virtual machine scheduling in private cloud faces many problems to achieve high performance and load balancing criteria. The hybrid energy-efficient scheduling approach helps to process the least time consuming process first so that the least scheduled process in queue need not wait for longer time that reduces the performance by increasing the turnaround time.

In paper[10] deals about the significant growth in size of the data to be stored in data centers is experienced, so this resulted in increased power consumption. For this reason managing of such energy resource is essential. Thus several architectures which can reduce power utilization in data centers is analyzed and studied.

2. Proposed Technique

This paper proposes a new algorithm called Enhanced Load Balancing algorithm (ELB) which maintains the global queue and implements part file method. The proposed algorithm provides an efficient solution for one of the challenging problems in VM scheduling namely the balancing of workload across compute nodes. The virtual machines are scheduled based on the tasks in the global queue onto the compute nodes primarily using classic round robin algorithm.

The round robin algorithm is a traditional way to balance workloads across compute nodes. Here each virtual machines does not maintain a separate queue instead a global queue is maintained which reduces the waiting time and increases the response time of the algorithm. In Load Balancing algorithm, the tasks are scheduled onto the virtual machines. The task which exceeds the queue length of particular virtual machine is taken as a whole task to the next virtual machine without utilizing the remaining memory space of the previous virtual machine. But this proposed algorithm gives a solution called a part file method, which partitions the task that exceeds the queue length, based on the accommodation of previous virtual machine. This part file method reduces the consumption of memory space in each virtual machine and increasing the efficiency of the algorithm.

3. ELB Algorithm

The proposed algorithm maintains an n task set from T1 to Tn. Let us assume that the arrival time for all the tasks in the task set is 0. The proposed algorithm implements part file method using Round Robin.

Input: Task set (T1, T2… Tn), Burst time (B1, B2… Bn)

- Initialize n = number of nodes, vm = number of virtual machines, q = queuelength, t = number of tasks. (as per Table 2).
- Obtain the inputs for burst time of all the task set.
- The tasks in the Global Queue (GQ) are scheduled to the corresponding Virtual machines (V).
- The task's burst time which exceeds virtual machine queue length is partitioned and the part file are equally scheduled to corresponding virtual machines.
- Until the end of respective node is reached.
- The finish time and turnaround time for each and every task in virtual machine are evaluated.

- The average turnaround time is calculated.
- The last task of the node which exceeds the queuelength is not partitioned and it is scheduled to the virtual machine of the next node.
- Repeat steps 4-8 until all the task in the GQ are serviced.

4. Experimental Analysis

In this paper, the proposed algorithm working is illustrated using a task set as referred in Table 1. The test unit comprises of four personal computers (DELL3147) each of which consists of 4GB RAM, 500GB HDD which is connected in a private cloud of an institution. It comprises of cloud scheduler and three compute nodes which supports up to four virtual machines. The process is controlled by the cloud scheduler. The tasks are scheduled onto the virtual machine in a circular fashion from the global queue.

In virtual machine the scheduling process happens as given below (Figure 1), suppose queuelength is assumed to be 20 then the task T1 and T2 is allocated in the first virtual machine. The next task T3 exceeds the queuelength so it is partitioned (T3 = 5 + 10). The first part of T3 (5) is scheduled to the first virtual machine and part file T3 (10) is transferred to the next virtual machine of the particular node.

Table 1. Task set and burst time as per paper[1]

Task Set(T)	Burst Time(B)
T1	5
T2	10
T3	15
T4	15
T5	5
T6	10
T7	20
T8	15
T9	5
T10	20

Table 2. Inputs and their description

Input	Description
N	Number of nodes
Vm	Number of virtual machines
Q	Queuelength
T	Number of tasks

5. Comparison Analysis

Turnaround time of task is the difference between the finish time and arrival time of a particular task. In comparison analysis the average turnaround time of Enhanced Load Balancing (proposed) and Load Balancing algorithm are compared in Table 3 and its corresponding figure is shown in Figure 2. The disadvantage of LB algorithm is average turnaround time is high and if a single task burst time exceeds the queuelength of the virtual machine then the entire task is not serviced. This is overcome by the part file method in the proposed algorithm (ELB).

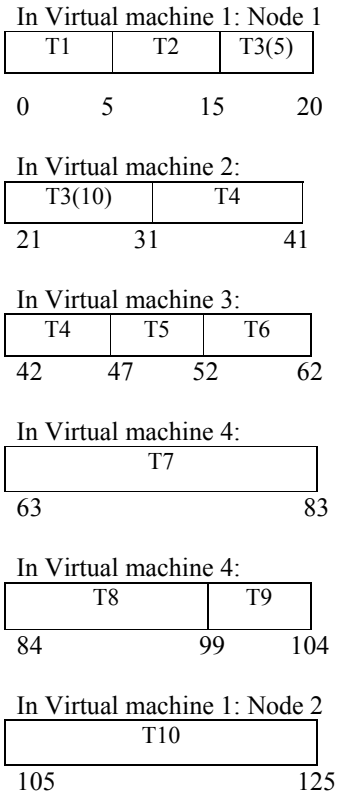

Figure 1. Proposed Gantt chart diagram.

Table 3. Comparison between ELB and LB

Algorithm	Average Turnaround Time
Load Balancing algorithm(LB)	8.35
Enhanced Load Balancing (ELB)Algorithm(proposed)	6.25

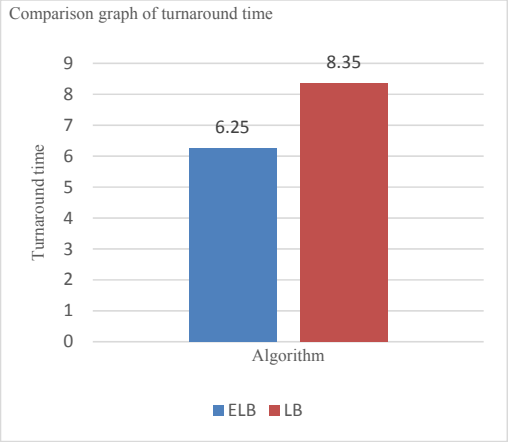

Figure 2. Comparison of turnaround time between ELB and LB.

6. Conclusion

The proposed algorithm uses round robin algorithm for scheduling the task to the virtual machine and implements part file method to improve the performance of the algorithm. From the experimental analysis and comparison analysis, ELB (proposed) algorithm produces high throughput with low turnaround time. ELB Algorithm implements a global queue that increases the response time and efficiency of the algorithm rather than maintaining a separate queue for each virtual machine. The part file method improves the usage of memory space in the virtual machines.

7. References

1. Li J. An energy-efficient scheduling approach based on private clouds. Journal of Information and Computational Science. 2011; 8(4):716–24.
2. Foster I. Cloud computing and grid computing 360-degree compared. Grid computing Environment workshop GCE'08. 2008. p. 1–10.
3. Li J, Zhang W, Peng J, Zhou L. A carbon 2.0 framework based on cloud computing. IEEE/ACIS 9th International Conference on Computer and Information Science; 2010. p. 153–8.
4. Lefurgy C, Rajamani K, Rawson F, Felter W. Energy management for commercial servers. Computer. 2003; 36(12):39–48.
5. Xu X, Cheng C, Xiong J. Reliable integrated model of cloud and client computing based on multi-agent. Journal of Computational Information Systems. 2010. p. 4767–74.
6. Round-Robin (RR). Available from: http://en.wikipedia.org/wiki/Round-robinscheduling
7. Armbrust M, et al. Above the clouds: A berkeley view of cloud computing. Berkley: University of California; 2009. p. 1–23.
8. Verma A, Ahuja A, Neogi A. Power-aware dynamic placement of HPC applications. Proc of the Annual International Conference on Super Computing, ICS'08; 2008. p. 175–84.
9. Verma A, Ahuja A, Neogi A. Power and migration cost aware application placement in virtualized systems. Proc of Middleware Conference; Leuven: Belgium. 2008. p. 243–64.
10. Bradley DJ, Harper R, Hunter S. Workload-based power management for parallel computer systems. IBM Journal of Research and Development. 2003; 47(5):703–18.