A task scheduling algorithm based on Qos

JunweiGe[1,a], Qingling Wang[2,b] and YiqiuFang[2,c]

[1,2,3]School of Computer Science and Technology, Chongqing University of Posts and Telecommunications, Chongqing, 400065, China.

[a]gejw@cqupt.edu.cn, [b]18996387624@163.com,[c] fangyq@cqupt.edu.cn

Keywords: Cloud computing, Task scheduling, Qos, User Satisfaction, Load balance.

Abstract. The key problem of task scheduling is to give optimal matching schemes for tasks and resources. In the scheduling process, users pursue high quality and service providers favor high yields,so the high efficiency must be considered in the process of task scheduling.The author puts forward a task scheduling algorithm on the premise of fully considering the system load-balancing, improving user satisfaction. By taking the simulation experiments of Cloudsim,The author proves that the algorithm is better than the algorithm of Min-Min based on Qos and the traditional algorithm of FCFS(First Come First Service).

1 Introduction

As a new business model,cloud computing is the product of integration and development of distributed computing, parallel computing, virtualization technology, grid computing and other technologies[1-2]. The task scheduling of cloud computing is the process of matching tasks with corresponding resources under the constrain of scheduling rules.If the mapping between tasks and resources is not reasonable, it can lead to the waste of resources and the serious influence to the utilization of system resources[3]. Thus, on the premise of meeting the requests of users and tasks, That how to completing tasks quickly and efficiently has become the emphasis of research.

2 Related work

In recent years, the task schedulingof cloud computing has attracted the people in all walks of life to study[3-4].In literature[5], the author considerd the minimization of taskcost from the perspective of service provider, buthe didn't consider the system load, which resulted in a waste of resources. In literature[6],the author proposed a load-balancing task scheduling algorithm which made a better balance of the load and reduced the completion time of the task.By increasing the optimal desired sequence and the index of load balancing, as well as using the number of task selections to constrain server connections, it achieved the goal of optimization. But it didn't consider the cost of the task, resulting in the high cost of execution. So this paper proposed a new task scheduling algorithmwhich considered the satisfaction of users and service providers,the load character of cloud system. Through the verification of the experment, this algorithm can improve degrees of satisfaction of tasks, reduce the completion time of tasks and improve the utilization of resources.

3 Related parameters defined ofthe process of task Scheduling

3.1 The definition of VM and task

VMis the acronym of the virtual machine.This paper made some assumptions as the following:

1)There is no interdependencies between different tasks ;

2)The load of task is related to the length of tasks and the speed of execution of the VM, the speed of execution of VMis related to the execution efficiency of the CPU.

Definition1: There is a quad of $(Id, Length, E_{price}, E_{time})$. Id stands for an unique identification of the task; $Length$ stands for the length of the task; E_{price}stands for the fee of the task that user desired,

E_{Time} stands for the completion time of the task that user desired.

Definition 2: The performance parameter of VM can be represented by (Id, Cpu, Bw, Mm, Hd). *Id, Cpu, Bw, Mm, Hd* respectively stands for the unique identification of VM, the task execution capabilities of the VM,the bandwidth of the VM, the virtual machine's memory size and the capacity of the VM's hard disk.

3.2 The Qos parameter Settings of task

In this paper, we consider two parameters which can be extended to more parameters.

1)The completion time of the task: users who chases for a short time to complete the task will choose this parameter.

2)The cost of the task: cloud computing is characterized by the pay-on-demand, so it is necessary to consider the cost of the task scheduling.

3.3 Parameters need to be considered in the process of task scheduling

Definition 3:Thecompletion timeof task *i* on VM *j* is defined as following:

$$Time_{ij} = Time_{run_{ij}}; (1)$$

Definition 4 : The cost of task *i* in virtual machine *j* is defined as following:

$$Cost_{ij} = (P_{Cpu_j} + P_{Bw_j} + P_{Mm_j}) * T_{r_{ij}}; (2)$$

P_{cpuj}, P_{Bwj}, P_{Mmj} respectively indicate the cost of each performance parameter on VM *j* in unit time.T_{rij} indicates the execution time of task*i* in VM *j*.

Definition 5: The satisfaction loss value of user is defined as following:

$$Uqos_{ij} = qos_{t_{ij}} + qos_{p_{ij}}; (3)$$

$$qos_{p_{ij}} = \begin{cases} \sqrt{\left(f(ct_{ij}) - E_{P_i}\right)^2}; & f(ct_{ij}) \geq E_{P_i} \\ 0; & f(ct_{ij}) < E_{P_i} \end{cases} (4) \quad qos_{t_{ij}} = \begin{cases} \sqrt{\left(f(T_{ij}) - E_{T_i}\right)^2}; & f(T_{ij}) \geq E_{T_i} \\ 0; & f(T_{ij}) < E_{T_i} \end{cases} (5)$$

qos_{pij}, qos_{tij} respectively indicates the satisfaction loss value of the execution time and the execution cost of the task. As the increase of function value,the satisfaction of users will decrease.

$f(ct_{ij})$, $f(T_{ij})$ stands forthe cost, the time of virtual machine *j* to complete task *i*. E_{pi},E_{Ti} indicatesthe fee, the expected timethat are users expected to pay for task *i*.

Definition 7 The definition of the load of VMas following:

$$Load_{ij} = Time_{wait_{ij}} / Load_j; (7)$$

$Load_{ij}$ stands for the ratio of the wait time of task *i*and the maximum loadof VM_j, the maximum load of VM_j defined as following: $Load_{\max_j} = Mm_j * HD_j / Cpu_j; (8)$

3.4 The description of Algorithm of BOQ(Based on Qos)

Before beginning, we should introduce three parameters:

1) $Load_{pt}$ stands for the preset thresholdof all VMs.

2) $Aqos_i$ representsthe sum of *Uqos* of all the tasks.

3) $Aqos\_loss_{ij}$ stands for the change valueof *Uqos* of the task when exchanged the scheduling position of task *i*and task *j*. The value of $Aqos\_loss_{ij}$ is decided by formulas (9), and $Uqos_{ij}$ stands for the satisfaction of task *m* when task *i*, *j* aren't exchange position.$Uqos_{ji}$ stands for the satisfaction of task *m*when task *i*, *j* exchange position.

$$Aqos\_loss_{ij} = \sum_{m=0}^{n} Uqos_{ij} - Uqos_{ji}; \qquad (9)$$

4)Ω_{TAll} stands for the set of tasks waiting to be distributed.

3.4.1 The specific flow of algorithm:

Algorithm:
Step1: Sorting*Uqos*of tasks in orderand mark the VM having the last element as VM_{most}.
Step2: Sorting the load of *VM*.
Step3: Using$Load_{VMqosmax}$to denote the load of *VMqosmax*.
The VM with smallest load will be marked as VM_{min}.

Step4: If task*i*can get the most *Uqos*in $VM_{most}$and the load of *VMqosmax*Less than $Load_{pt}$, tie task *i* and *VMj* directly and delete the task from Ω_{TAll} and schedule the next task. Otherwise, execute the step 5.
Step5: Iterating all tasks in the waiting queue of VM. If the $length_i$is shorter than$length_m$(*m* inVM_{most}'swaiting queue), exchange their position. Calculate the $Aqos\_loss_{im}$ after exchanging. If the $length_i$is more than $length_m$, not exchange them and execute them according to the initial distribution.
Step6: Sorting all $Aqos\_loss_{ij}$, and marking the largest one as $Aqos\_loss_{ijmax}$.
Step7: If $Aqos\_loss_{ijmax}<0$, we should not exchange them and decrease the value of$Aqos_i$. If $Aqos\_loss_{ijmax}>0$, we should exchange them , so $Aqos_i$will increase.
Step8: Returning Ω_{TAll}. Deleting the task*i*, then executing step 1 until Ω_{TAll} is empty.

4 Simulation and analysis of experiments

In order to verify the validity of this algorithm, the parameter of task are setted are as follows: the range of the number of tasks is [0, 250]. The tasks length range from 1 to 40000. The Qos Sequence of expectation is [Q_{Price}, Q_{Time}]. The value of $Q_{Price}$and Q_{Time} range from 0 to 1. And add $Q_{Time}$to Q_{Price}, the value is 1.

The parameters of VM involved in the experiment are shown in Table 1:

Table 1 parameter list of virtual machine

VM_{id}	Cpu	Bw	HD	MM
1	1024	800	800000	1024
2	512	1000	160000	1024
3	1024	1200	320000	2048
4	1024	700	600000	512

The experimental results as shown in Fig. 1 and Fig. 2.

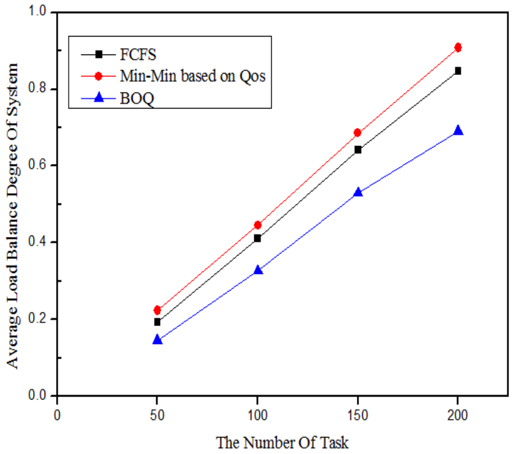

Fig. 1 The Comparison of System Average Load

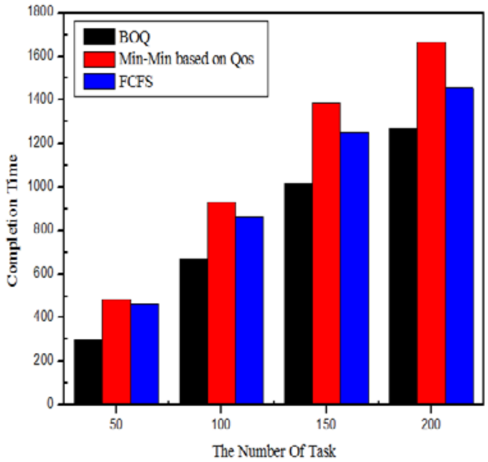

Fig. 2 The Comparison of completion time

From figure 1, figure 2, the algorithm proposed in this paper is superior to the Min - Min algorithm based on Qos and FCFS algorithm on the degree of load balancing and task completion time.

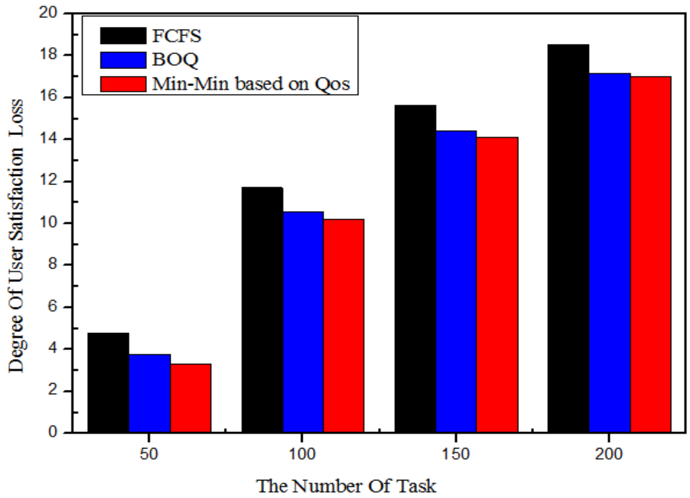

Fig. 3 The Comparison of degree of user satisfacton loss

In Figure 3, the degree of user satisfaction loss of the algorithm of BOQ is not better than the Min - Min algorithm based on Qos. Considering from three aspects, the reason for the effect of the algorithm of BOQ is that the Min - Min algorithm based on Qos and FCFS only focus on the profits of single aspects and the algorithm of BOQ pays attention to system load-balancing when it pursues high user satisfaction.

5 Conclusion

The task scheduling algorithm has proposed in this paper not only can keep the load balance of the system, it improve the utilization of system resources by using the dynamic load balancing strategy. By using the method of switch position of tasks, it can reduce the satisfaction loss value of task and improve the user satisfaction of batch task, but the algorithm still has room for improvement, it needs to be studied further.

References

[1]Liang Guicai. Research Survey of Cloud Computing[J]. computer application, 2014, S2:70-72+77.

[2]Venters W, Whitley E A. A critical review of cloud computing: researching desires and realities[J]. Journal of Information Technology, 2012, 27(3): 179-197.

[3]edersen J M, Riaz M T, Dubalski B; et al. Using latency as a QoS indicator for a global cloud computing services[J]. Concurrency and Computation Practice and Experience, 2013, 25(18): 2488-2500.

[4]Hirai T, Masuyama H, Kasahara S, et al. Performance analysis of large-scale parallel-distributed processing with backup tasks for cloud computing[J]. Journal of Industrial and Management Optimization, 2014, 10(1): 113-129.

[5]Deng Jianguang. Research on Task Scheduling Strategy of Cloud Computing[D]. South China University of Technology, 2014.

[6]Zhang La, Liu Shufen, Han Lu. Algorithm for Tasks Scheduling Based on Load Balance[J]. Journal of Jilin University(Science Edition), 2014, 04:769-772.