Research of cloud computing task scheduling algorithm based on improved genetic algorithm

GE Junwei

College of Computer Science and Technology
Chongqing University of posts and
Telecommunications
Chongqing, China
gejw@cqupt.edu.cn

YUAN Yongsheng

College of Computer Science and Technology
Chongqing University of posts and
Telecommunications
Chongqing, China
yys8128394@163.com

Abstract—**Use genetic algorithm for task allocation and scheduling has get more and more scholars' attention. How to reasonable use of computing resources make the total and average time of complete the task shorter and cost smaller is an important issue. The paper presents a genetic algorithm consider total task completion time, average task completion time and cost constraint. Compared with algorithm that only consider cost constraint (CGA) and adaptive algorithm that only consider total task completion time by the simulation experiment. Experimental results show that this algorithm is a more effective task scheduling algorithm in the cloud computing environment.**

Keywords-cloud computing; genetic algorithm; task scheduling

I. INTRODUCTION

Clouding computing [1] has become a focus for researchers in recent years. There are many companies involved in cloud computing, and obtain the results. Such as Google's cloud computing platform, IBM's "Blue Cloud" platform product and Amazon's elastic compute cloud. and so on.

Cloud computing is a new computing model developed on the basis of parallel computing, distributed computing, cluster computing, grid computing, besides have the all the features of above calculation mode, it also has following advantages: on-demand, efficient, reliable, low-cost. It will put huge computational processing program automatically split into numerous smaller subroutines through the network, then give the vast system which is composition of a number of servers to calculate, the processing result after calculate and analysis will back to the user. According to the type of service, it can be divided into three categories: IaaS, consider infrastructure as a service, PaaS, consider the platform as a service, and the SaaS, consider software as a service.

II. SCHEDULING PROBLEM ANALYSIS

In the cloud computing environment, how to schedule a large number of tasks is a complicated problem [2], Cloud computing need to provide services to multiple users at the same time, and should be consider the response time and the average completion time of each user, also taking into account the cost of service [3-4]. Some of the existing scheduling algorithms usually only consider optimize one of

the factors, which may cause the task completion time is short, but the cost is relatively high. Therefore, I propose a scheduling algorithm (MGA), which comprehensive consider total task completion time and average task completion time and cost for the task, it improved task scheduling and resource allocation strategy for the cloud computing, so as to maximize the efficiency of cloud computing.

III. IMPROVED GENETIC ALGORITHM

Genetic Algorithm (GA) [5] is relatively complete theory and method by Professor Holland and his colleagues, students research and formation, from trying to explain the natural systems of biological complexity adaptive process proceed with, simulate biology evolution mechanism to construct artificial system model. They trying to explain biological complexity adaptive process in the natural systems, then simulate biology evolution mechanism to construct artificial system model. Its essence is a kind of high efficient, parallel, the global searching method. It can automatic acquisition and accumulation of knowledge of the search space in the search process, and adaptive control the search process in order to achieve the optimal solution [6].

A. chromosome encoding and decoding

Provided m task, namely, t_1, t_2, t_3..., the number of worker resources is m, the i task (t_i) has been divided into n subtasks, then the j-th subtask of i-th task marked as $t_i(j)$, task number in total:

$$num = \sum_{i=1}^{m} \sum_{k=1}^{n} t_i(j) \qquad (1)$$

Assuming there are 3 tasks, 3 worker resources, t_1 divide into 5 sub task: $t_1(1)$, $t_1(2)$, $t_1(3)$, $t_1(4)$, $t_1(5)$; t_2 divide into 5 sub task $t_2(1)$, $t_2(2)$; t_3 divide into 3 sub task $t_3(1)$, $t_3(2)$, $t_3(3)$;there are 10 sub tasks in total, encoding the 10 sub tasks according to the task order, length of chromosome is 10, each value of genes is between 1 and 3, generate the chromosomes are as follows:

{3,2,1,1,1,2,2,2,3,1}

The chromosome represents $t_1(1)$ execute in third worker, $t_1(2)$ execute in second worker, $t_1(3)$ execute in first worker, and so on.

After decoding it to obtain the distribution and execution sequence of subtasks on each resource, the chromosome decoding for:

W_1：$\{ t_1(3) , t_1(4) , t_1(5), t_3(3)\}$

W_2：$\{ t_1(2), t_2(1), t_2(2), t_3(1)\}$

W_3：$\{ t_1(1) , t_3(2)\}$

Though decoding, it can get sub task sequence of each worker needed to complete, then the use ETC matrix can calculated complete time of assigned task by for each worker:

$$\text{workertime}(k) = \sum_{j=1}^{n} time\,(k,j), k\epsilon[1,w] \qquad (2)$$

time(k, j) represents a k-th worker on the time required to complete the j-th task.

Because worker can be executed in parallel, so the total task completion time is the longest completion time of workers, that means, the total task completion time is:

$$\text{sumtime} = \max_{k=1}^{w} worktime(k) \qquad (3)$$

Time is required by the i-th task completion:

$$\text{tasktime(i)} = max_{k=1}^{w} \sum_{j=1}^{s} time(k,j) \qquad (4)$$

s is the location of subtask of task i assign to the worker, The execution time of task i subtasks on the worker is $\sum_{j=1}^{s} time(k,j)$, w is worker number.

Average task completion time:

$$\text{averagetime} = \frac{\sum_{i=1}^{m} \text{tasktime(i)}}{m} \qquad (5)$$

Cost is required in the sub-tasks are completed:

$$\text{cost} = \sum_{k=1}^{w} \text{workertime}(k) \times RCU(k) \qquad (6)$$

RCU (k) represents the cost of each the computing resource unit time task

B. Production of the initial population

Assuming S is the scale of population, T is the total number of tasks, each task is divide into several subtask, M is total number of subtask, W is the resource number, also is the worker number. Then the initialization described as: generated S chromosomes by randomly, M is chromosome length, Gene ranges [1, W].

C. Fitness function

Genetic algorithm is based on some fitness function for the next generation of choice, thus to find the optimal solution of the problem. Therefore the fitness selection is very important, which is related to the convergent speed of the algorithm and the quality of obtained solution.

In the task scheduling, an important goal is total task completion time is short, but the average time needed for the task cannot be ignored, while considering the service costs, thus defining a three fitness function.

The fitness functions of total task completion time:

$$f_1(t) = \frac{1}{\text{sumtime(t)}} , 1 \le t \le S \qquad (7)$$

sumtime(t) is the all subtasks completion time of the t population

The fitness functions of average task completion time:

$$f_2(t) = \frac{m}{\sum_{i=1}^{m} tasktime(i,t)} \qquad (8)$$

$\sum_{i=1}^{m} tasktime(i,t)$ is the task average completion time of t population

The fitness functions of cost:

$$f_3(t) = \frac{1}{cost(t)} \qquad (9)$$

cost(t) is required cost of the t population.

The fitness functions:

$$f(t) = \alpha f_1(t) + \beta f_2(t) + \gamma f_3(t) \qquad (10)$$

Here, $\alpha\epsilon[0,1], \beta\epsilon[0,1], \gamma\epsilon[0,1], \alpha + \beta + \gamma = 1$.

When $\alpha+ \beta=0, \gamma=1$, it means fitness function only consider Cost constraints. When $\alpha+ \beta=1, \gamma=0$, it means fitness function only consider total completion time and average time of task constraints.

D. genetic operations

1) Selection operation

Selection is used to determine the recombination or crossover individual, and the being selected individual will generate how many offspring individuals. Firstly, calculate the fitness, then is the actual selection, Selecting the parent individual in according with the fitness. This paper adopts the roulette wheel selection as the selection operator, it is similar to roulette in the gambling game, it can be calculated the probability of individual will be selected according to the fitness function:

$$P(t) = \frac{f(t)}{\sum_{j=1}^{s} f(j)} \qquad (11)$$

2) Crossover and Mutation

Cross is the operation which put portions structure of two father individual replace and recombinant to generate new individual. The searching ability of genetic algorithm improves a lot by the cross. Crossover operator will random exchange certain gene between two individuals based on cross ratio in the population, which able to produce a new gene combinations and expected useful genes together. Mutation operator's basic content is changes genetic value of some Gene locus in the Individual string among the population, which can improved the local random search capability of genetic algorithm and maintain the population diversity.

In this paper, the crossover probability function and mutation probability function as follows:

$$P_c = \begin{cases} P_{c1} - \frac{(P_{c1} - P_{c2})(f' - f_{avg})}{f_{max} - f_{avg}}, f' \ge f_{avg} \\ P_{c1} \qquad\qquad\qquad , f' < f_{avg} \end{cases} \qquad (12)$$

$$P_m = \begin{cases} P_{m1} - \frac{(P_{m1} - P_{m2})(f_{max} - f)}{f_{max} - f_{avg}}, f \ge f_{avg} \\ P_{m1} \qquad\qquad\qquad , f < f_{avg} \end{cases} \qquad (13)$$

Here, $P_{c1} = 0.9$，$P_{c2} = 0.6$，$P_{m1} = 0.1$，$P_{m2} = 0.001$, f_{max} is the largest fitness value in the population, f_{avg} is the average fitness value of each generation population, f' is the larger fitness value in the cross two individuals, f is the fitness value of mutation individual.

IV. ALGORITHM SIMULATION AND RESULTS ANALYSIS

This paper uses the CloudSim[7] platform to simulate the cloud computing environment, in the same environment, using CGA and AGA algorithm with MGA do the comparison test.

The initial conditions of the algorithm: the value of S is 50, the value of W is 50, the value of T is 50, and the value of M is 1000. ETC matrix and a RCU array are randomly generated by the system.

The termination condition of the algorithm: set the maximum generations is 100, the curve tends to be stable at this time, so think the algorithm basic convergence, terminating algorithm

The parameters in the algorithm as shown in Table 1:

TABLE 1 EXPERIMENTAL PARAMETERS

Parameters	Algorithm		
	MGA	CGA	AGA
α	0.5	0	1
β	0.2	0	0
γ	0.3	1	0
P_{c1}	0.9		
P_{c2}	0.6		
P_{m1}	0.1		
P_{m2}	0.001		

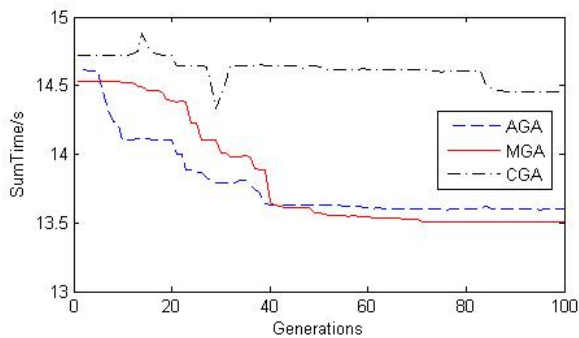The experimental results are shown below:

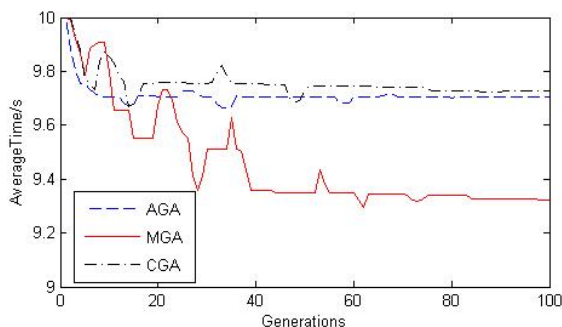Figure 1.　Total task completion time comparison
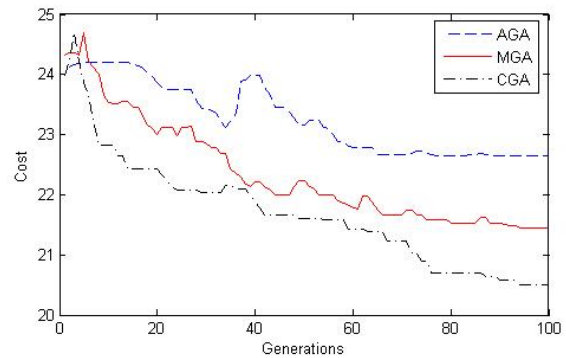

Figure 2.　Average task completion time comparison

Figure 3.　Cost Comparison

From Figure 1 and Figure 2, in the evolution initial stage of genetic algorithm，both the total task completion time and the average task completion time of AGA are slightly lower than total task completion time and the average task completion time of MGA. However, with the growth of the Generations, Because of AGA only concerned about the total task completion time previously, resulting in the loss of some potential good genes and fall into the local convergence; The MGA with the growth of Generations, the total task completion time and average task completion time were less than the total task completion time and the average task completion time of AGA; But total task completion time and the average task completion time of CGA was not obviously optimized.

It can be seen from Figure 3 that there is little difference between the cost of the three required early in the evolution. But with the growth of Generations, the MGA and CGA cost is getting smaller and smaller, while the cost of AGA was no obvious optimization.

V. CONCLUSIONS

This paper presents a the genetic algorithm (MGA) which comprehensive consideration of the total task completion time, average task completion time and cost constraints, it can obtained not only total task completion time and average task completion time is short, but also the required costs is low. The result of this task scheduling is ideal.

REFERENCES

[1] CHEN Quan, DENG Qianni, "Cloud computing and its key techniques", Journal of Computer Applications, vol. 29, no. 9, pp.2563-2566, Sep9. 2009.

[2] WU Yuqi, ZENG Guosun, ZENG Yuan, "Trend Analysis for Scheduling Algorithm in Cloud Computing", Journal of Microelectronics & Computer, vol. 29, no. 9, pp.103-107, Sep9. 2012.

[3] R Buyya, D Abramson, J Giddy, "An economy driven resource management architecture for global computational power grids. Int'l Conf on Parallel and Distributed processing Techniques and Applications, Las Vegas, 2000

[4] XU Xiaoyong, PAN Yu, "Power-aware resource scheduling under cloud computing environment", Journal of Computer Applications vol. 32, no. 7, pp.1913-1915, July1. 2012.

[5] WANG Xiaoping; CAO Liming, Genetic Algorithms, Xi'an Jiaotong University Press, China, 2002.

[6] Vincenzo Di Martino. Scheduling in a grid computing environment using genetic algorithms. Marco Mililotti the 16th Int'l parallel and Distributed processing Symp(I PDPS2002),Florida ,USA,2002.

[7] Rodrigo N. Calheiros, Rajiv Ranjan, César A. F. De Rose, Rajkumar Buyya, "CloudSim: A Novel Framework for Modeling and Simulation of Cloud Computing Infrastructures and Services",2009