International Journal of Computer Networks and Communications Security

VOL. 3, NO. 6, JUNE 2015, 253–258
Available online at: www.ijcncs.org
E-ISSN 2308-9830 (Online) / ISSN 2410-0595 (Print)

Optimal Scheduling In Cloud Computing Environment Using the Bee Algorithm

Msc.NASRIN HESABIAN[1, 2] and phD.HAMID HAJ SEYYED JAVADI[3]

[1] Department of Computer Engineering, Science and Research Branch, Islamic Azad University, Boroujerd, Iran

[2] Department of Computer Engineering, Boroujerd Branch, Islamic Azad University, Boroujerd, Iran

[3] Department of Applied Mathematics, Faculty of Mathematics and Computer Science, Shahed University, Tehran, Iran

E-mail: [1, 2]n.hesabian@gmail.com, [3]hamid.h.s.javadi@gmail.com

ABSTRACT

Cloud computing has made a fundamental change in the way of strange information and data and implementation of application progress. Everything is hosted on a cloud that is a set of several servers and computer, which can be accessed through internet instead of placing data and application programs on a personal computer. The challenge of cloud computing system is dedicating the resources to the system requests. Dedicating resources to the requests is a NP-complete problem due to requests and resource dynamics. In recent year, one of the most important and promising method to solve such problems is innovatin methods inspired from the nature. These methods are similar to the social or natural system. In this article, we want to use honeybee colony algorithm for resources scheduling. This algorithm is an optimization method based on swarm intelligence and intelligent behavior of honeybee population. Honeybee algorithm involves a group based on search algorithm.

Keywords: *Cloud Computing, Scheduling, Resource Dedication, Honeybee Algorithm.*

1 INTRODUCTION

Cloud computing is a general and newfound technology delivering resources and application program to the users. Multi-national companies such as Microsoft, IBM, Google and Oracel companies provide various services of cloud computing to the clients. Currently, cloud computing is a commercial field [1],[2]. In terms of cloud computing, companies organizations and individuals don't pay money for software, hardware or the network; instead, they buy the required software services and computing power. This idea follows high saving and productivity in information technology resources. In fact, cloud computing is a pattern of distributed computing , and it is a combination of many resources and requests with the aim of resource subscription in the form of a service in the internet. The challenge of cloud computing systems is dedicating the resource to the system requests. Dedicating the resource to requests is a NP-complete problem due to requests and resources dynamics [3],[4]. in cloud computing systems, computing resource are presented as virtual machines . In such scenario, scheduling algorithms play a very important role because the aim is scheduling the efficiency of tasks so that time is reduced, and resource exploration can be improved. In recent years, one of the most important and promising methods to solve such problems is innovative methods inspired from the nature. These methods are similar to social and natural systems. These methods are genetic algorithm, ant-colony optimization algorithm and bee-colony optimization problem. In this article, we want to use bee colony algorithm in resource scheduling. This algorithm is an optimization

method based on swarm intelligence and intelligent behavior of bee population. Since bee algorithm is a meta-innovative method based on distributed systems and parallel processing technique, it is used for scheduling optimization [15].

2 CLOUD COMPUTING

Distributed systems and parallel processing techniques are the solution for better and quicker utilization of complex information in the present. Nowadays, there are hundreds of computers and supercomputers with various architecture capacities in all over the world, and they are used in scientific, military and commercial fields. Mostly, information subscription among them is necessary. A distributed system as a set of independent computers and users consider it as a coherent system, grid computing is a distributed system and it is infrastructure of cloud computing. This technology provides accessing remote resources by using computer networks and communication infrastructure as well as by using concepts and facilities of distributed systems. Computing resource of heterogeneous software and hardware systems can be connected to each other without geographical limitations, and whole system structure seems as a unique and integrated virtual machine. Then, very large and complex application, program, requiring high processing and much input data can be implemented in this virtual machine. In fact, the purpose is to use computing resources of systems to perform tasks when they are free [5],[6].

Cloud computing has grid properties as well as under properties. In fact, cloud computing as a pattern of distributed computing, and it is combination of many resources and request with the aim of resource subscription in the internet [7, 8].

3 LITERATURE REVIEW

Since tasks scheduling is dynamic in cloud computing environment, it is difficult to obtain a proper solution that can minimize tasks implementation, the time of executing input tasks and can consider load balance between resources.

The number of request that has been entered to the system is more and higher than the number of resource, and it does not stop [9],[10].

* The properties of requests that home been entered to the environment are unclear, and they continuously change such as entering time, executing time and required memory.

* Cloud computing environment is a set of heterogynous resources.

* Resources home some hardware and software properties such as load volume, the amount of free memory in a system, it has properties of communication network such as bandwidth, traffic and etc. Kanlee and his colleagues (2011) proposed scheduling of cloud tasks on the basis of load balancing ant colony optimization algorithm (LBACO)[17]. The main policy of this method is load balance of whole system. Also, it has been attempted to minimize the time of Make span. In this article, the properties of ant algorithm has been considered to schedule the tasks. Sourav Banerjee and his colleagues (2002) presented scheduling algorithm based on genetic algorithm[11].

Kennedy and his colleagues (2001) proposed an algorithm that has been inspired from the behavior of birds and fish , and it is in unline discrete routs used in ants colony algorithm. This algorithm called particles swarm algorithm searches the solution space by adjusting various factors.

Babu and his colleagues (2003) used bee meta-heuristic algorithm to reach load balance in machines. The proposed algorithm provides tasks priority in machines queue on the basis of the least waiting time. They found considerable improvement in execution time and reducing waiting time in queue [12][13][14][16].

4 THE PROPOSED ALGORITHM

As we know, various users send various tasks to the system, so different computing capacities are required. Dedicating the capacity, similar to cloud resources, to various users is undesirable. In addition, it is assumed that cloud resources are heterogeneous, independent and dynamic. Hence, resources properties are temporal and they change. Cloud performance depends on scheduling method because cloud environment is dynamic, independent and heterogeneous, and the space of tasks is complicated.

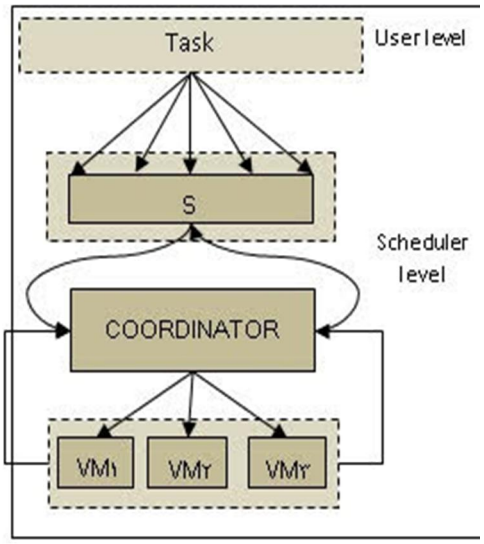

Fig. 1. the assumptions of cloud system for proposed algorithm.

For example, as it has been shown in figure1, the cloud system involves scheduling of S and three cloud nodes of VM={VM1,VM2,VM3}. In cloud system, we consider S= {VM1, VM2, VM3}. As it has been already said, the number of cloud node is equal to the number of schedulers. Bees involve three groups in bee colony algorithm. They involve employed bees, supervision. They involve employed bees, supervisory bees and scout bees. The bee staying in dancing area to make decision and select the food source is called explorer bee. The bee going toward predetermined food source is called employed bee, while the bee performing random search is called scout bee. The main steps of algorithms are as follow:

- Initial initialization
- Repetition

A. The location of employed bees in food sources in the memory
B. The location of search bees in food sources in the memory
C. Sending scout bees to search new food sources

- (The considered situation is obtained).

In ABC algorithm, each search cycle is composed of three stages.

These stages are sending employed bees toward food sources and then measuring the amount of their nectar, selecting food sources they explorer bees after subscribing information they employed bees, determining the amount of nectar, determining scout bees and then sending them to food sources. In the stage of initial initialization, a set of food sources are randomly selected by the bees and the amount of their nectar is detected. Then, these bees come to the hive, and data and informa-tion of nectar in each hive is subscribed with the bees waiting in dancing area. In the second stage, after information subscription, each employed bee goes to the range of food source this bee has already visited this range, and has stored that food source in its own memory. Then, a new food source is selected by using visual information existing in the neighborhood. In the third stage, a supervisory bee select the food source depending on nectar information and data distributed by the employed bees in dancing area. The nectar amount of food source increases, and the food source selected by supervisory bee increases. Therefore, dancing employed bees carring much nectar encourage supervisory bees toward food source with much nectar. The bee selects a new food source in the neighborhood depending on visual information after entering to the selected area. Visual information is on the basis of comparing food source. When the nectar of food source is releasd by the bees, a new food source is randomly determined by the scout bee. The algorithm is performed by a set of solutions, and it‘s attempted to improve it called population.

In this article, some bees have been considered as scout bees, and some have been taken into account as supervisory bees, and some of them are considered as employed bees. After returning to the hive and after dancing other bees, they go to the neighborhood of the places where other bees have moved through observing their dance. Two methods have been used to simulate it. A random number is created. If this random number is lesser than threshold limit, then, the first node occurs. If it is more than threshold limit, the second mode occurs. It should be mentioned that the set of tasks and virtual machines are defined as follows:

VM= {VM1, VM2,...., VM3}
Task= {T1, T2,...Tn}

Each bee is the answer of a problem, so the bee is shown as follows:

Bee=[X11, X21... Xij]

Xij shows the ith task processing in jth processor. For instance, X12 shows that the first task is processed in the second processor. Therefore, in order to simplify it, the bee has been shown as follows:

B=2 3 1 2 1 3

This bee shows that the first task is dedicated to the second processor, and the second task is dedicated to the third processor and so on. After creating initial bees or population, the fitness of each bee is computed by using fitness function. In order to compute the fitness, two parameters involving loading balance and makespan of the last task are simultaneously used as follows:
K1: impact factor of makespan of the last task
K2: impact factor of loading balance

$$F([B_i]) = k_1 \times Makespan([B_i]) + k_2 \times \frac{1}{Load\ .Balancing([B_i])\%}$$

F([Bi]) shows fitness of bee i. [Bi] demonstrates the values of bee Bi. With regard to above definition, the smaller the amount of F, the higher the fitness will be. After computing the fitness of each bee, some of them are selected as scout bee. Other bees move to the neighborhood of the places where these bees have gone. In order to compute the neighborhood location, the method of defining threshold limit is used the third stages. Some bees are randomly sent to the environment again, so that placing them in the local optimally is prevented. After creating all bees, bee's population is updated. Generally, the stages are as follows:

1. Creating the initial population
2. Computing the fitness function
3. Ordering the population based on fitness function.
4. Defining the threshold limit

- If the random number is lesser than threshold limit, then some locations are randomly selected. Random numbers are created by a uniform distribution between [-1, 1]. The selected locations created by random numbers are summed up.

- If the random number is lesser than threshold limit, nodes information is extracted from coordinator. A specified percentage of nodes having higher volume of processing is determined. A specified percentage of the nodes having the least volume of processing is determined. The load of nodes in moved.

5. Computing fitness function
6. Selecting the best bees
7. Producing the new initial population
8. If the number of replications is enough, it ends; otherwise, return to stage 3.

5 EVALUATION AND COMPARISON

In order to compare the proposed method with other methods, three parameter involving makespan, flow time and the average of waiting time.

In this article, we want to compare the results obtained from our proposed method (PM) with the results obtained from the proposed algorithm (dhinesh BABu, 2013)[16]. Using honey meta-heuristic algorithm to reach the load balance over the machines, it is called HBB-LB.

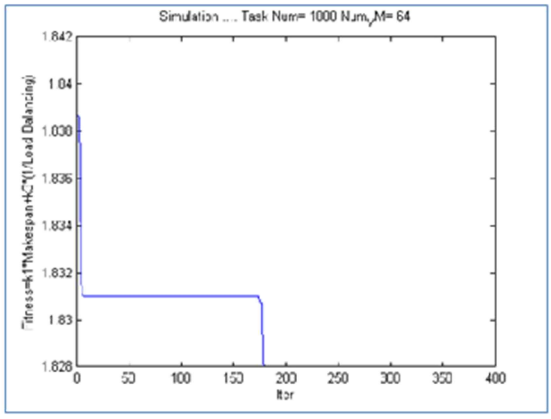

Fig. 2. changes of fitness function with the number of replication (1000 tasks-64 virtual machines)

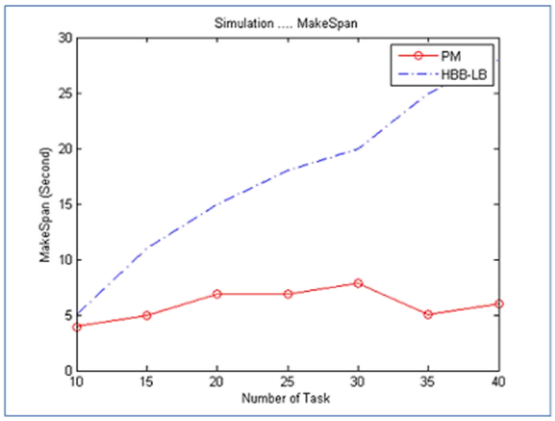

Fig. 3. comparing the flow time in the proposed method

Table 1 shows the changes of makespan values before and after executing the proposed algorithm for 1000 tasks and 64 resource, as it has been shown, by applying the algorithm , the results obtained from makespan reduces from 54 seconds to 18 seconds; that is, the reduction of 66% is observed.

Table1: comparing various parameters after and before applying the proposed method.

Parameters	Before	After	Percent
MakeSpan	54.86	18.17	66.87
Load Balancing	68	90	32.35
Mean Wait Time	1.93	0.92	52.33
Degree of Imbalance	1.84	0.5	72.82

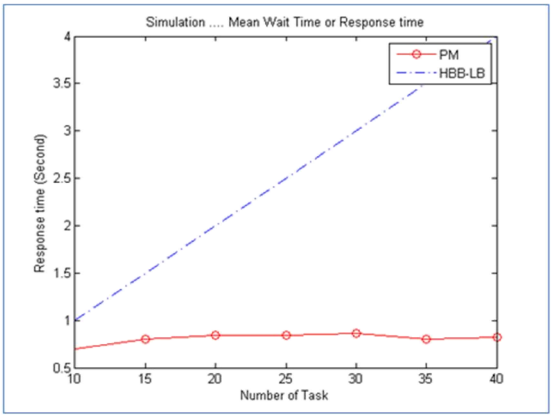

Fig. 4. comparing the average of waiting time in the proposed method.

Figure 4 demonstrate the average of waiting time in the method proposed in the method in comparison to HBB-LB method .Like makespan , responding time dose not reduce by increasing cloud environment scale. HBB-LB method not only bees weaker performance than PM but also its performance reduce by increasing the network scale. Response time is one of the parameters, showing the service quality in the network. Therefore QOS of the proposed method is better than HBB-LB method.

6 CONCLUSION

As we know, various users send various tasks to the system, so different computing capacity is required. Dedicating capacity similar to cloud resources to various users is undesirable. In addition, it's assumed that cloud resources are heterogeneous, independent and dynamic; therefore, the propertice of cloud resources are temporary, and they change. Cloud performance depends on scheduling method since cloud environment is dynamic, independent and heterogeneous and the space of tasks is complex.

In this article, we presented a scheduling method in cloud computing environment on the basis of bee algorithm to dedicate the sources optimally. The proposed method has been compared with other algorithms such as HBB-LB. the proposed method or PM method behaves like HBB-LB method in small environment in terms of makespan. Its performance becomes better by increasing the environment dimensions. It can be due to the nodes awareness of each other's position. Their task is to cordinate these tasks.

7 REFERENCES

[1] Arkaitz ruiz-Alvarez, Marty Humphery, A Model and Decision Prosedure for Date Storage in Cloud Computing, in Proceedings of the IEEE/ACM International Symposium on Cluster, Ottawa Canada, 2012.

[2] S.Selvarani, Dr.G.Sudha Sadhasivam , "IMPROVED COST-BASED ALGORITHM FOR TASK SCHEDULING IN CLOUD COMPUTING" in IEEE, India 2010 .

[3] Bhaskar Prasad Rimal, Eunmi Choi, Ian Lamb, " A Taxonomy and Survey of Cloud Computing Systems" in IEEE, Fifth International

[4] B. Basturk ; D. Karaboga, "An artificial bee colony(abc) algorithm for numeric function optimization," in IEEE Swarm Intelligence Symposium 2006, Indianapolis,Indiana, USA, May 2006.

[5] D. Karaboga , B. Basturk; " On the performance of artificial bee colony (ABC) algorithm" in Elsevier Erciyes University, Engineering Faculty, Computer Engineering Department, TR-38039 Kayseri, Turkey 2008.

[6] Rajkumar Buyya ,Chee Shin Yeo, Srikumar Venugopal, James Broberg, Ivona Brandic, "Cloud com puting and emerging IT platforms: vision, hype, and reality for delivering computing as the 5t h utility ",Future Generation Computer systems, ELSEVIER (2008).

[7] Steve Bennett, Mans Bhuller, Robert Covington ," Architectural Strategies for Cloud computing " , August (2009) .

[8] Brucker P., Drexl A., Mohring R., Neumann K., Pesch E .,"Resource-constrained roject scheduling: Notation, classification, models, and methods", Elsevier European Journal of Operational Research, Vol. 112, pp. 3-41, 1999

[9] Malathi G, Sarumathi S. 2010. Survey on grid scheduling.

[10] Alcaraz J., Maroto C.,"A Robust Genetic Algorithm for Resource Allocation in Project Scheduling", Annals of Operations Research, Vol. 102, pp. 83-109, 2001

[11] Brucker P., Knust S., Schoo A., Thiele O.,"A branch and bound algorithm for the resource-

constrained project scheduling problem", Elsevier European Journal of Operational Research, Vol. 107, pp. 272-288, 1998

[12] Delavar A.G., Rahmany M., Halaakouie A., Sookhtsaraei ,R "DSQGG: An optimized Genetic-based algorithm for Scheduling in Distributed Grid ",2nd International Conference , Computer Technology and Development (ICCTD), (2010).

[13] Pham D.T., Ghanbarzadeh A., Koc E., Otri S., RahimS.,M.Zaidi, "The Bees lgorithm - A Novel Toolfor Complex Optimisation Problems," Proceedings of IPROMS 2006 Conference, pp. 454-461, 2006.

[14] D. Karaboga ; B. Basturk, "A powerful and efficient algorithm for numerical function optimizationArtificial bee colony (abc) algorithm," Journal of Global Optimization, p. 459–471, 2007.

[15] Sourav Banerjee ,Mainak Adhikari ,Utpal Biswas "Advanced Task Scheduling for Cloud Service Provider Using Genetic Algorithm" IOSR Journal of Engineering (IOSRJEN) ISSN: 2250-3021 Volume 2, Issue 7(July 2012), PP 141-147

[16] Dhinesh Babu L.D, P. Venkata Krishna "Honey bee behavior inspired load balancing of tasks in cloud computing environments" journal home page: www.elsevier.com/l ocate/asoc (Applied Soft Computing 13 (2013) 2292–2303) , 2013 Elsevier .

[17] Kun Li, Gaochao Xu, Guangyu Zhao, Yushuang Dong, Dan Wang "Cloud Task scheduling based on Load Balancing Ant Colony Optimization" 2011 Sixth Annual ChinaGrid Conference, 978-0-7695-4472-4/11 $26.00 © 2011 IEEE.