2010 2nd International Conference on Industrial Mechatronics and Automation

A Load Balancing Mechanism Based on Ant Colony and Complex Network Theory in Open Cloud Computing Federation

Zehua Zhang and Xuejie Zhang
School of Information Science and Engineering, Yunnan University, Kunming, P.R. China
{zehuazh, xjzhang}@ynu.edu.cn

Abstract—**Although cloud computing is generally recognized as a technology which will has a significant impact on IT in the future. However, Cloud computing is still in its infancy, many crucial problems need to be solved for the realization of the fine scenery which theoretically depicted by cloud computing. Load balancing is one of these problems; it plays a very important role in the realization of Open Cloud Computing Federation. We proposal a load balancing mechanism based on ant colony and complex network theory in open cloud computing federation in this paper, it improves many aspects of the related Ant Colony algorithms which proposed to realize load balancing in distributed system, Furthermore, this mechanism take the characteristic of Complex Network into consideration. Finally, the performance of this mechanism is qualitatively analyzed, and a prototype is developed to enable the quantitative analysis, simulation results manifest the analysis.**

Keywords- load balancing; ant colony; complex network; cloud computing; federation

I. INTRODUCTION

With the quick development and extensively use of IT technology, cloud computing become a hot topic of industry and academe as an emerging new computing mechanism, with a great hope that it will provide computing as the 5th utility (after water, electricity, gas, and telephony) to meet the everyday needs of the general community [1], Cloud computing represents a distributing computing mechanism that by the utilize of the high speed network, data processing is moved from private PC or servers to the remote computer clusters (big data centers owned by the cloud service providers).

The conception of Open Cloud Computing Federation (OCCF) is proposed by many researchers from industry and academe [2, 3]; it incorporates multiple CCSP's (Cloud Computing Service Provider) service to provide a uniform resource interface for the user. We think that OCCF is the only way to get the cloud computing be widely used and realize the greatest value of it, since the main value of cloud computing is brought by the scale economy effect which is caused by the share and reuse of resources by lots of users.

Despite the glorious future of Could Computing, many crucial problems still need to be solved for the realization of cloud computing. Load balancing is one of these problems, it plays a very important role in the realization of Cloud Computing and Open Cloud Computing Federation, the function of load balancing is to distribute the excess local workload evenly to the whole cloud federation, aims to realize a high ratio of user satisfaction and facility utilization in the cloud. Load balancing is indispensable for cloud computing. Firstly, the CCSP must use load balancing in its own cloud computing platform to provide a solution with high efficient for the user; Secondly, a inter CCSP load balancing mechanism is needed to construct a low cost and infinite resource pool for the consumer.

The scale of a cloud computing platform would be very large, and provides various kinds of operation systems for the consumer, and applications running on the cloud are numerous. Such a cloud environment can be treated as a complex system which has the characteristic of dynamic, open, non-linear interact and great amount of components. Load balancing in the cloud should be distributed, flexible and extensible. Although many works has been done on the problem of load balancing in different kinds of distributed system (such as grid, telecommunication and internet) with the ant colony optimization technology, there is still room for improvement, and the complex network's characteristic isn't be considered in these works.

In our previous work [4], we presented a MABOCCF (Mobile Agent Based Open Cloud Computing Federation) model which realizes portability and interoperability between different Cloud Computing platforms, as a result of one of the mobile agent's ability that can overcome heterogeneous. Based on MABOCCF, we proposal a load balancing mechanism based on ant colony and complex network theory in open cloud computing federation in this paper, it improves many aspects of the related Ant Colony algorithms which are proposed to realize load balancing in distributed system, Furthermore, this mechanism take the characteristic (small-world and scale-free) of Complex Network into consideration, thus a better performance on load balancing is gained.

The rest of this paper is organized as follows: Section 2 gives a review of the related work which realizes load balancing with the ant colony optimization technology. Section 3 introduces the characteristic of Complex Network. Section 4 describes the load balancing mechanism based on ant colony and complex network theory in open cloud computing federation. Section 5 analyses the performance of the mechanism. Finally, a conclusion of the work is discussed in Section 6 along with the envisaged future work.

II. RELATED WORK

Individual ants in some ways are much unsophisticated insects. They have a very limited memory and exhibit individual behaviour that appears to have a large random component. Acting as a collective, however, ants manage to perform a variety of complicated tasks with great reliability and consistency. A colony of ants can collectively perform useful tasks such as building nests, and foraging (searching for food) [5, 6]. Drawing upon some of the computing techniques inspired by social insects such as ants [5], several mobile agent-based paradigms were designed to solve load balancing problems in distributed systems.

Montresor et al have developed an inverse artificial ants system, where artificial ants disperse a group of tasks evenly on idle nodes [7]. In the SearchMax process, an ant wanders across the network to look for an overloaded Node, then, in the SearchMin process, the ant wanders across the network to look for an underloaded node, finally, the ant transfers a task from the most overloaded node to the most underloaded one. J. Cao's work appoints a fixed move step m to the SearchMax process and the SearchMin process in [8]. [7, 8] are most related to the work described in this paper, but the convergence speed can be enhanced if the characteristic of the network's structure is considered in their work.

An agent-based load balancing mechanism in homogeneous minigrid environments is presented in [9]. Each ant represents a task, so, the agent-based load balancing is regarded as agent distribution from a macroscopic point of view. An ant's strategies in moving between nodes are decided by the number and size of teams where agents (tasks) queue. Although this mechanism is a convenient model for quantitative analysis, it's a classic ant colony optimization (ACO) algorithm.

In [10], an ant-like mechanism is presented to achieve load balancing in telecommunications networks, however, we think that the move of the mobile agents can be further improved in this mechanism.

III. CHARACTERISTIC OF COMPLEX NETWORK

Many systems in nature and in technology consist of a large number of highly interconnected dynamical units. Despite the inherent differences, most of the real networks are characterized by the same topological properties, as for instance relatively small characteristic path lengths, high clustering coefficients, etc. All these features make real networks radically different from regular lattices and random graphs, the standard models studied in mathematical graph theory. We only list the small-world property and scale-free distribution here [11].

Small-world: In most of the real networks, despite of their often large size, there is a relatively short path between any two nodes. This feature is known as the small-world property and is mathematically characterized by an average shortest path length L, defined as in (1), which depends at most logarithmically on the network size N.

$$L=\frac{1}{N(N-1)}\sum_{i,j\in N,i\neq j} d_{ij} \qquad (1)$$

Where d_{ij} is the length of the geodesic from node i to node j.

Scale-free distribution: When the scientists approached the study of real networks from the available databases, it was considered reasonable to find degree distributions localized around an average value, with a well-defined average of quadratic fluctuations. In contrast with all the expectancies, it was found that most of the real networks display power law shaped degree distribution as in (2), with exponents varying in the range $2<\gamma<3$.

$$P(k) \sim A^{-\gamma} \qquad (2)$$

These two characteristics are considered for the move of the ants in our work, since the node with a large degree may leads the ants (or the tasks) move more quickly towards the region where more resources may be found for the execution of the tasks.

IV. A LOAD BALANCING MECHANISM BASED ON ANT COLONY AND COMPLEX NETWORK THEORY IN OPEN CLOUD COMPUTING FEDERATION

After more then a decade of study, mobile agent has get a lot of breakthroughs in many key technologies, but it is suffered from that it can't find a appropriate platform in a large scale network to deploy and manifest its expected advantages. Cloud computing technology provide a good a chance for mobile agent to reveal its capability, we think cloud is the best platform for mobile agent so far. Mobile agent based cloud computing is feasible because most of mobile agent systems are based on or support Java, and many cloud computing platforms currently provide virtue machines which support different kinds of OSs (operating systems) such as Linux and windows, since Java can "write once, run anywhere", so the mobile agents can run on the JVMs (Java Virtual Machine) install on these OSs. In our previous work [4], we presented a new mechanism call MABOCCF (Mobile Agent Based Open Cloud Computing Federation) which combine mobile agent with cloud computing to construct a open cloud computing federation mechanism, with hopes to realize portability and interoperability between different kind of cloud computing platforms. Based on MABOCCF, we proposal a load balancing mechanism based on ant colony and complex network theory in open cloud computing federation for OCCF in this paper.

A. The components and the initialization of the complex network

A open cloud computing federation (OCCF) is consist of many cloud computing service provider's (CCSP) facilities, there would be many management regions (partitioned by geographically or the management strategies) of one CCSP, a node in each management region is chosen as the region load balancing node (RLBN), each RLBN connect with many of the other RLBNs of a CCSP according to the information get from the CCSP, after that, many of a CCSP's RLBN is selected to connect with RLBNs in other CCSP, so the topology of the connected RLBNs makes a overlay network which can be regarded as a complex network, the structure of the OCCF can be depicted as in fig.1.

A load can add to or remove from the management region; also, a RLBN can be elected from the region nodes once a RLBN is failed. Afore mentioned construction behaves is the initialization state of the complex network, the evaluation of the complex network caused by the ant is given in IV.B.

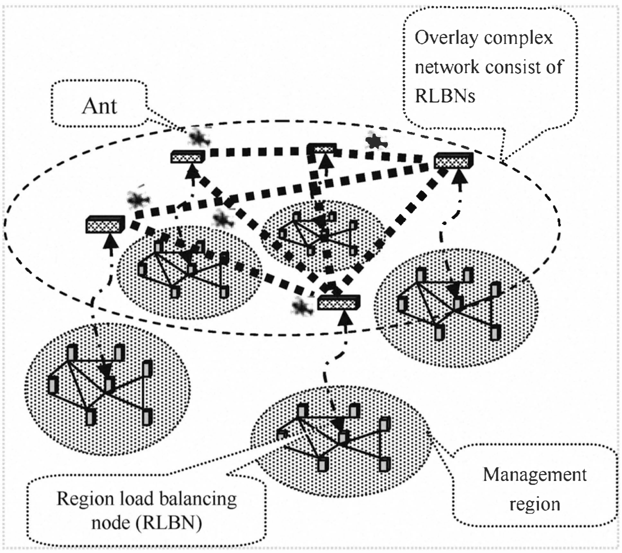

Figure 1. The structure of the OCCF and the formation of the complex network

B. Realization of load balancing with ant algorithm

1) Underload load balancing method: An ant is periodically sent out by the underload node to balance the workload on the whole OCCF (open cloud computing federation), and keep the complex network's vitality by update the pheromone on each node through the following steps.

a) Once an ant start its trip from a node, at each move of the ant, the probability function for an ant at node N_i to choose a neighbor node N_j as its next stop at time t is:

$$P_{ij}(t)=\begin{cases}\dfrac{\left[\tau_{ij}(t)\right]^{\alpha}\left[\eta_{ij}(t)\right]^{\beta}\left[D_{ij}(t)\right]^{\lambda}}{\sum_{j\in N_n}\left[\tau_{ij}(t)\right]^{\alpha}\left[\eta_{ij}(t)\right]^{\beta}\left[D_{ij}(t)\right]^{\lambda}} & if \quad j\in N_n \\ 0 & if \quad j\in N_n\end{cases} \tag{3}$$

Where N_n is N_i's neighbor nodes set which haven't been visited by the ant, τ_{ij} is the pheromone value at edge d_{ij}, $\eta_{ij}=1/d_{ij}$ is the cost of edge (i, j), d_{ij} is the length of the edge (i, j), and D_{ij} is the proportion of N_j's degree to the total degree of N_i's neighbor nodes, D_{ij} can be calculated as:

$$D_{ij}=\frac{d_j}{\sum_{k\in NB_i} d_k} \tag{4}$$

Where NB_i is the set of neighbor nodes of Ni. This means the nodes with a big degree has more opportunity to be chosen as the ant's next stop. α, β and λ represent the respective adjustable weights of τ_{ij}, η_{ij} and D_{ij}.

b) During its trip, the ant remember the node which has max/min workload (noted is N_{max} and N_{min}, respectively) and the corresponding workload on them, the ant stop at a certain node when the difference of the workload between N_{max} and N_{min} exceeds a threshold T or the total move steps of ant attains a previous appointed number m. Then, the ant informs the two nodes N_{max} and N_{min} to balance the workload between them.

2) Overload load balancing method: Once a node find its workload has excess its own threshold W, an ant is sent out by the node, then, the following processes is like that in the underload load balancing method except that the ant's source node is appointed as the N_{max} in this method.

Many hypo-max or hypo-min nodes can also be remembered by the ant to improve the load balancing performance of this mechanism.

3) The update of the pheromone

Once the load balancing is performed between the nodes N_{max} and N_{min}, the ant backtrack the path it has traversed to update the pheromone values on the edges (each node maintains a pheromone table of the edges which link to its neighbor nodes, the values corresponding to the amount of pheromone on the edge is stored in this table), in our mechanism, the data in the table, and the data of the proportion of the pheromone changed on each edge to the pheromone changed on the whole path (including the increment and the evaporation) are all normalized, and the distribution of the pheromone changed on the whole path is previously assigned according to the strategies that more pheromone is assigned to the edges near the N_{max} and less pheromone to the edges near N_{min}.

4) The evolution of the complex network

In our mechanism, the structure of the complex network should evolve to adapt to the change of work load distribution, after the update of the pheromone, the addition of a new edge between the nodes N_{max} and N_{min} is considered if there is not such a edge and the Cost Effectiveness is greater than a threshold F_{max}. On the contrary, the deletion of an existing edge is considered if it "sleeping" too long. Therefore, a complex network with the characteristic of small-world and scale-free is expected to be gained through the local behaviors of the swarms of the ants, theses two characteristic is useful in the load balancing processes of the ant algorithm, in other words, the node with a large degree in such a complex network may leads the ants move more quickly towards the region where more resources may be found for the execution of the tasks.

V. PERFORMANCE ANALYSIS

A open cloud computing federation (OCCF) is consist of many cloud computing service provider's (CCSP) facilities, it contains numerous of applications which comprise millions of modules, these applications serve for large quantities of users, moreover, the resource and the requirements of user is dynamic. In such situations, how to get a high user satisfaction and a better facility utilization ratio is a really a complex problem. To meet with this require with this requirements, we proposal a load balancing mechanism based on ant colony and complex network theory in open cloud computing federation for OCCF in this paper, the advantages of this mechanism is described as fallows.

1) Overcome heterogeneous of the cloud computing facilities. Currently, the reality is that the API of different cloud computing platform is private, each CCSP has poured a large amount of funds on their own cloud computing facilities, it's impossible for them to abandon current platform and adopt a new cloud computing standard. A mobile agent based load balancing system can run on the heterogeneous facilities with the support of java and virtue machines, little work is done and with modification on the existing facilities.

2) Adaptive to the dynamic environments in the OCCF. The swarms of ants can behavior adaptively according to the workload variation on different management nodes and the change of the degree distribution of the nodes in the overlay network, this lead to a good performance in load balancing speed.

3) Excellent in fault tolerance. This mechanism is a completely distributed system and with no central control point.

4) Good scalability. Any resource can randomly add in the OCCF with little effect on the system performance, and the load balancing can be quickly gain under the abnormal emergence situations of the heavy workload.

We construct a prototype with Java to compare our work (abbreviated as ACCLB) with the other two mechanisms. One is the SearchMax and SearchMin mechanism (abbreviated as SMM) presented in [7, 8], another one is a classic ant colony mechanism (abbreviated as CAC) like in [10]. Fig.2 depicts the overall standard deviation of the workload data obtained from the simulation after 5 to 50 iterations of the three mechanisms with the increment of 5 iterations. In a dynamic environment with a complex network structure (which is the most cases), the overall standard deviation decreases more quick than the other two mechanisms, and a more evenly distribution of the workload on the whole cloud federation is gained, the simulation results confirmed our analysis. The results were obtained in a simulation of a complex network of 100 nodes that generated by a java algorithm in advance. Initially, 10,000 tasks of the same size are generated and distributed evenly on 10 nodes which are randomly selected. The nodes in the complex network have the same Execute capability *EC*=100workload/iteration, another 10000 tasks of workload is added randomly to ten of the nodes in the complex network after each iteration step.

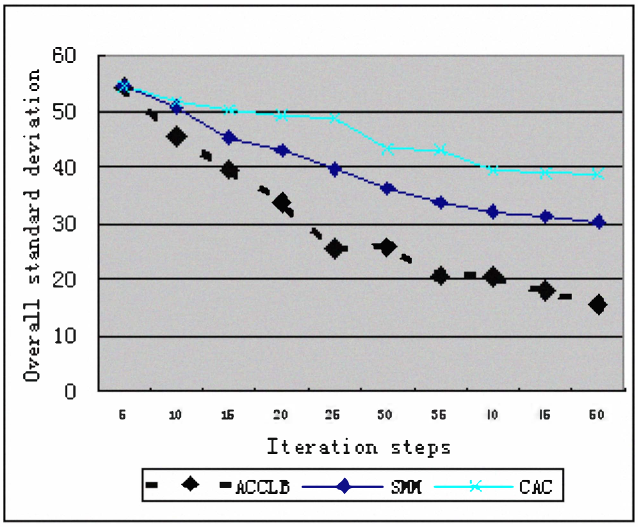

Figure 2. Simulation result of ACCLB and SMM

VI. CONCLUSION AND FUTURE WORK

A open cloud computing federation (OCCF) is consist of many cloud computing service provider's (CCSP) facilities, the goal of load balancing is to distribute the excess local workload evenly to the whole cloud federation, and realize a high ratio of user satisfaction and facility utilization in the cloud. We proposal a load balancing mechanism based on ant colony and complex network theory for OCCF in this paper, which aims to cope with the complex and dynamic load balancing problem in OCCF. This mechanism improves many aspects of the related Ant Colony algorithms which proposed to realize load balancing in distributed system, and the characteristic (small-world and scale-free) of Complex Network have been taken into consideration.

Many study shown that the dynamic change of the parameter in the calculation of the probability function for an ant to choose a neighbor node may bring a good performance to the ant colony optimization (ACO) algorithm. How the parameters α, β and λ in (1) effect each other will be further studied in our future work.

ACKNOWLEDGMENT

This work is supported by the National Natural Science Foundation of China (NSFC) (No.60573104) and the Natural Science Foundation of Education Department of Yunnan Province (No.09Y0050).

REFERENCES

[1] R. Buyya, C. S. Yeo, S. Venugopal, J. Broberg, I. Brandic, "Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility," Future Generation Computer Systems, vol.25, pp. 599-616, June 2009.

[2] B. Rochwerger, D. Breitgand, E. Levy, A. Galis, K. Nagin, I. Llorente , et al. "The Reservoir model and architecture for open federated cloud computing," IBM Journal of Research and Development, Volume 53, April 2009.

[3] R.Ranjan, R.Buyya, "Decentralized Overlay for Federation of Enterprise Clouds," http://arxiv.org/ftp/arxiv/papers/0811/0811.2563.pdf.

[4] Z Zhang, X Zhang, " Realization of open cloud computing federation based on mobile agent," in Proc. Of IEEE International Conference on Intelligent Computing and Intelligent Systems (ICIS 2009), Shanghai, China, pp.642-646, 2009

[5] E.Bonabeau, M.Dorigo, and G.Theraulaz, "Inspiration for optimization from social insect behavior," Nature, vol.406, pp.39–42, July2000.

[6] M.Dorigo, G.D.Caro, and L.M.Gambardella, "Ant algorithms for discrete optimization," Artif.Life, vol.5, no.2, pp.137–172, 1999.

[7] A. Montresor, H. Meling, and Ö. Babaoglu, "Messor: Load-Balancing through a Swarm of Autonomous Agents", in Proc. of 1st Int. Workshop on Agents and Peer-to-Peer Computing, 1st ACM Int. Joint Conf. on Autonomous Agents and Multi-Agent Systems, Bologna, Italy, 2002.

[8] J. Cao Self-Organizing Agents for Grid Load Balancing In Proc. of 5th IEEE/ACM International Workshop on Grid Computing, Pittsburgh, PA, USA, pp. 388-395, 2004.

[9] J. Liu, X. Jin and Y. Wang, "Agent-Based Load Balancing on homogeneous Minigrids: Macroscopic Modeling and Characterization," IEEE Transactions on Parallel and Distributed Systems, Volume 16, NO.6, JUNE 2005.

[10] R. Schoonderwoerd, O. Holland, and J. Bruten, "Ant-like agents for load balancing in telecommunications networks," in Proc. Agents, Marina del Rey, CA, USA, pp.209–216, 1997.

[11] S. Boccalettia, V. Latorab, Y. Morenod, M.Chavezf, D. -U. Hwanga, "Complex networks: Structure and dynamics," Physics Reports, 424, pp.175 – 308, 2006.