Optik 127 (2016) 5099–5104

Contents lists available at ScienceDirect

Optik

journal homepage: www.elsevier.de/ijleo

Virtual machine resource scheduling algorithm for cloud computing based on auction mechanism

Weiwei Kong [a,b,*], Yang Lei [c], Jing Ma [d]

[a] Xijing University, Xi'an 710123, China
[b] Department of Information Engineering, Engineering University of Armed Police Force, Xi'an 710086, China
[c] Department of Electronics Technology, Engineering University of Armed Police Force, Xi'an 710086, China
[d] Key Laboratory of Information Assurance Technology, Beijing 100072, China

A R T I C L E I N F O

Article history:
Received 18 January 2016
Accepted 28 February 2016

Keywords:
Cloud computing
Virtual machine
Auction mechanism
Queuing theory

A B S T R A C T

To overcome the problem of virtual machine (VM) scheduling in cloud computing, a novel adaptive VM resource scheduling algorithm based on auction mechanism is presented by considering multiple factors including network bandwidth and auction deadline. First, the sequencing of the clients' bids is conducted in the given competition deadline. Second, the client group is screened and corresponding VM resource is configured according to the minimum costs of the cloud service providers. Finally, the final payment price can be figured by considering the levels of average payments and competitive payments, so that the tasks clients request can be completed with the given VM resource. The simulation experimental results show that the proposed algorithm can effectively enhance the quality of service of the cloud environment, the profits of cloud service providers and the resource utilization rate of VM.

1. Introduction

Cloud computing [1–3] is the integration product between the traditional computing model such as distributed computing, virtualization and the developing network technology. Based on the above computing model, the computing, storage, platform and service resource pools can be realized, which are abstract, dynamic extending and manageable. Moreover, the resource pools can be provided to external clients on demand via Internet.

Virtual machine (VM) is one of the most typical resources in cloud computing. The resource scheduling and allocation not only are related to quality of service (QoS) [4–6], but also can affect the profits of the cloud service provider directly. At present, resource scheduling has become the hot issue in the field of cloud computing, and a series of resource scheduling algorithms [7–21] have been presented by the scholars both at home and abroad.

Commonly, the scheduling issue boils down to certain mathematical problem first. Then, the corresponding mathematical model can be constructed. Finally, the solution to the related model is figured out. Traditional cloud computing algorithms include first-in-first-out (FIFO) scheduling algorithm, greedy algorithm, fair scheduling [7,8] algorithm, and so on. However, these algorithms belong to the static type, thus there is no adaptive and dynamic adjustment mechanism. Unfortunately, the cloud computing resources are usually dynamically allocated and released, therefore, traditional algorithms cannot satisfy the practical requirements of the cloud computing resource scheduling, and the resources are wasted seriously. There is a great deal of research that shows that resource scheduling is not only a multi-constrained multi-objective optimizing issue in essence, but also a NP problem. To this end, many artificial intelligence algorithms are introduced into the solution to the cloud computing resource scheduling, such as genetic algorithm (GA), particle swarm optimization [9], clone selection algorithm, shuffled frog leaping (SFL) [10] algorithm, and so on. Due to the high efficiency of the newly proposed algorithms, they have been the mainstreamed research basis. Ref. [11] integrated GA into the scheduling model, and the original fitness function was improved via QoS, but the result easily trapped in a local optimum. Yang et al. [12] proposed a novel heuristic optimizing algorithm called bat algorithm (BA) in 2010, and a variety of improved versions were devised to deal with the resource scheduling of cloud computing. Similarly, SFL model and its modifications were also utilized to conduct the resource scheduling.

* Corresponding author at: Xijing University, Xi'an 710123, China. Tel.: +86 15609231541.
E-mail address: kwwking@163.com (W. Kong).

http://dx.doi.org/10.1016/j.ijleo.2016.02.061
0030-4026/

Ref. [13] proposed a resource evaluation model based on entropy optimization and dynamic weighting. He et al. [14] presented a management model called elastic application container (EAC) to deal with the lightweight resource scheduling. In Ref. [15], an algorithm was devised to estimate the quantity of VM resources the service provider offers according to the demands of the clients require. Gong et al. [16] evaluated the VM resource demands based on the theories of both the signal and statistics. In Ref. [17], the differential calculation method was used to simulate the optimal dynamic VM resource sets. Experimental results demonstrated that the performance of the algorithm is good, but its computational costs were high. Different from most ones, Refs. [18–21] all involved the cloud resource distributing algorithms based on the auction mechanism. However, the algorithm in [18] needs the step of prior prediction on the workloads. In Ref. [19], we didn't have to conduct the prior prediction, but we paid for it with high computational complexities. By combining the superiorities of [18,19], Refs. [20,21] took many factors such as VM resources and network bandwidth into consideration, and improved the traditional auction mechanism a lot. However, due to the ignorance of the factor of competition deadline, the wasting of VM resources may appear to some extent.

Based on the background mentioned above, this paper proposes an adaptive VM resource scheduling algorithm which considers both auction mechanism and many factors such as network bandwidth and auction deadline. This algorithm consists of two phases. The one is the cloud property evaluation and VM configuration, and the other one is the client auction payment mechanism. The former is responsible for the sequencing of the clients' bids in the given competition deadline, so that the potential client group can be determined. The latter can realize the payment for the network resources and figure out the final payment price of the clients. The simulations of the algorithms involved in this paper are conducted on the CloudSim platform, and the experimental results demonstrate that the proposed algorithm can effectively enhance QoS of the cloud environment, the profits of the cloud service providers and the resource utilization rate of VM.

The rest of the paper is organized as follows. In Section 2, we give the traditional cloud computing model and four dynamic resource description indexes. The auction mechanism based VM resource scheduling algorithm and its concrete steps are given in Section 3. In Section 4, experimental results and related analysis are described. Finally, Section 5 draws a conclusion for this paper.

2. Traditional cloud computing model

Cloud computing can provide large batches of task requests for a large number of clients simultaneously. Once receiving the service requests, cloud service providers will distribute corresponding computing resources based on different requests from the clients or the dispenses of the cloud computing resources the clients pay for. The traditional cloud computing model is shown in Fig. 1.

The cloud computing resource especially the VM resource performance is an important index in the cloud computing service, which could measure the service probability and the level of service ability of the cloud service providers. Usually, VM resource performance evaluation is mainly based on the information of the VM resource available. The description of the current used and available VM resources can be realized by the evaluation of VM resource availability, which is beneficial for resource scheduling, distribution and transferring.

In order to efficiently describe the information of dynamic resources in the cloud computing, four dynamic resource description indexes are proposed in this paper, including request index (RI), dealing index (DI), loading index (LI) and cost index (CI).

Fig. 1. Traditional cloud computing model.

Suppose the VM resource set denoted by $\boldsymbol{S} = \{S_1, S_2, \ldots, S_n\}$, and n is the number of VM resources.

(1) *RI*: RI is the service request quantity the cloud service providers receive in unit time. The larger the RI value is, the more resource quantity of the cloud service providers the clients request. If the service request the VM resource $S_i (1 \le i \le N)$ receives in unit time is $\mathrm{RI}_i (1 \le i \le N)$, then the RI of the set $\boldsymbol{S}$ is:

$$\mathrm{RI} = \sum_{i=1}^{n} \mathrm{RI}_i \tag{1}$$

(2) *DI*: DI is the service request quantity VM resources accomplish in unit time. The value of DI reflects the dealing ability to the clients' requests of the cloud service resources. The larger DI is, the higher the dealing ability is. If the service request quantity the VM resource $S_i (1 \le i \le N)$ accomplishes in unit time is $\mathrm{DI}_i (1 \le i \le N)$, then the DI of the set $\boldsymbol{S}$ is:

$$\mathrm{DI} = \sum_{i=1}^{n} \mathrm{DI}_i \tag{2}$$

(3) *LI*: LI is the ratio between the average time required to complete a client's service for the VM resource and the clients' adjacent service request interval time. Similar to the above two indexes, the value of LI is proportional to the loading of the cloud service resources. With regard to the VM resource $S_i (1 \le i \le N)$, if the average time required to complete a client's service is $t_i (1 \le i \le N)$ and the clients' adjacent service request interval time is t_d, then the LI is:

$$\mathrm{LI} = \frac{t_i}{t_d} \tag{3}$$

(4) *CI*: CI is used to weigh the cost required to maintain the normal running of the VM resource. If the running maintenance cost of the VM resource $S_i (1 \le i \le N)$ is C_{iR} and the idle maintenance cost of the VM resource $S_i (1 \le i \le N)$ is C_{iI}, then the total maintenance CI of the cloud service provider is in the following range:

$$\mathrm{CI} = \left[\sum_{i=1}^{n} \mathrm{C}_{iI}, \sum_{i=1}^{n} \mathrm{C}_{iR} \right] \tag{4}$$