Advance Journal of Food Science and Technology 11(4): 281-285, 2016
DOI:10.19026/ajfst.11.2411
ISSN: 2042-4868; e-ISSN: 2042-4876

Submitted: May 5, 2015 Accepted: June 22, 2015 Published: June 05, 2016

Research Article

Scheduling Methods for Food Resource Management under the Environment of Cloud

Yang Jianhui
School of Mathematics and Statistics, Zhoukou Normal University, Zhoukou, China

Abstract: The core problem of cloud computing is to deal with the dynamic scheduling problem of physical and virtual resource, which can be regarded as a new problem in cloud computing, in order to achieve energy saving and emission reduction with high performance and low investment as well as some other targets. In this study, it takes the overview of the cloud computing technology as the cutting point, with the help of the interpretation of computing food resource scheduling under the environment of cloud, discussing the implementation of cloud computing task scheduling with optimization method of genetic algorithm.

Keywords: Cloud computing, food resource, scheduling problem

INTRODUCTION

With the development of cloud computing in various areas, there are a variety of problems occurred such as food resource management (Ye, 2011), food resource scheduling and security mechanism and so on, which are reflected day by day, while the strategy of scheduling virtual food resource has become the core issue of cloud computing research (Tan *et al*., 2011; Mark *et al*., 2011). The essence of scheduling food resource is to find out the appropriate resources and make reasonable distribution for a large number of requests, so as to meet the request of task, optimize time and improve the utilization rate of resources as far as possible (Bezdek and Pal, 1992).

MATERIALS AND METHODS

Overview of cloud computing technology: Cloud computing is a new business model and service model, its core concept is to be no longer dependent on the local computer to do the calculations, but it can be operated by the third party to have calculations and store resources (Tseng and Tzeng, 2002). It can distribute the computing tasks into the different data centers that are made up of a large number of physical servers and virtual servers, so as to make all kinds of application obtain computing power, storage space and information services according to the needs (Geng *et al*., 2013).

Cloud computing can integrate various large server cluster into a huge pool of resources to form as cloud, which can provide services through the Internet and make the underlying infrastructure abstract so that researchers can focus on their own areas, who does not need to care about other less important links. The service model of cloud computing can be shown in Fig. 1, including Software as a Service, Platform as a Service and Infrastructure as a Service (Li *et al*., 2015).

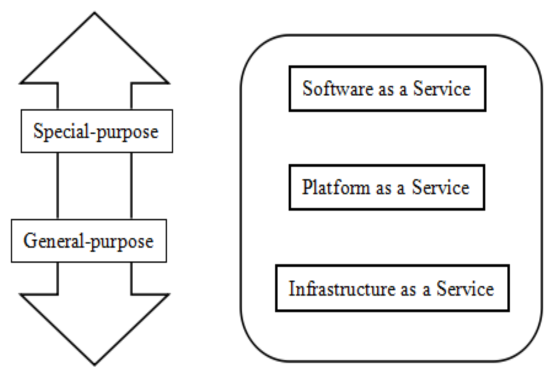

Fig. 1: Model of cloud computing service

Food resource scheduling in cloud computing environment: Food resource scheduling refers to the process that applied the available food resource to the requested task with food resource assignment, so as to ensure the efficiency of finishing task and keep the load of food resource balanced with the premise of completing tasks. Cloud computing is evolved on the base of grid computing technology, its food resource scheduling method can be referred to grid computing in some extent, but the different scenarios make them have significant difference between the two scheduling methods: grid computing is mainly used in the field of scientific research, which put attention to the efficiency, optimizing the user's requests for the scheduling

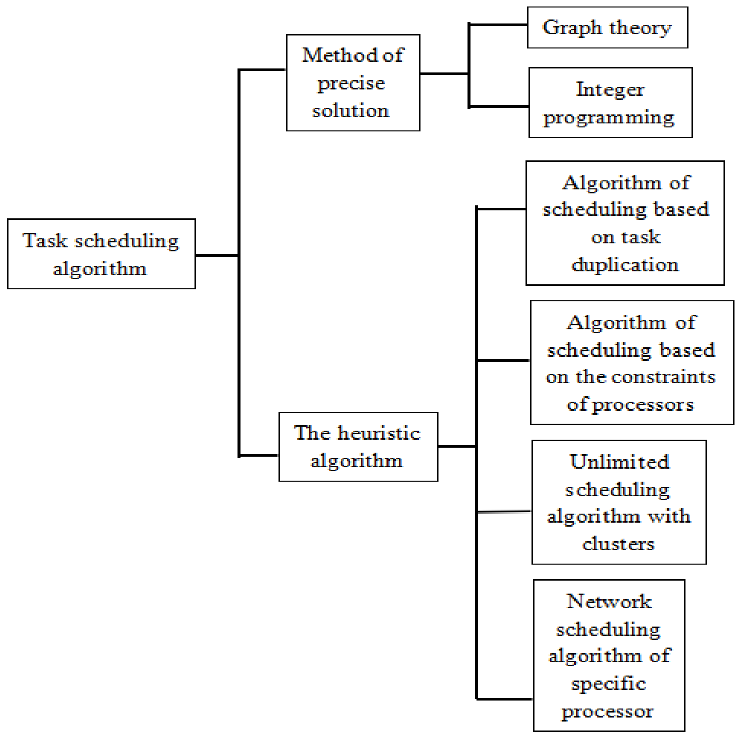

Fig. 2: Classification of task scheduling algorithm

objective, while cloud computing has to face a large number of users' needs and different applications and different application requests have different demands on quality of service, therefore, the required resources are also different. In addition, the types of application to cloud computing platform resources are widely distributed with many kinds of factors, which have increased the difficulty to manage resource.

During the process of food resource scheduling, as for a large number of food resource requests, firstly, it should detect the performance of virtual machine resources, computing the occupancy rate of current resources of the virtual resources as well as the allocation standard deviation function of the resource, real-timely understanding the load of virtual machine, trying to find out more suitable virtual resources to perform tasks, so as to reduce the load of resources and make the correspond request for resource.

Food resource scheduling algorithm is the core to realize the optimization of scheduling resource, which is the key of all kinds of resources and service systems that can provide users with the appropriate resources.

Combined with the previous research in the field of cloud computing, it can draw the following conclusions: the scheduling algorithm of Heterogeneous Distributed Computing (HDC) system can be divided into two main categories: one category is through looking for the given target to achieve the optimal food resource allocation method based on the exact solution, such as scheduling method of integer programming theory, however, the food resource optimization scheduling problem is a NP-Hard problem, for finding out the exact solutions need to pay the large cost of computing. Therefore, another kind used heuristic algorithm to find out the approximate solution of the method, which has been applied widely. The classification of task scheduling algorithm in calculation system can be shown in Fig. 2.

Critical scheduling calculations: In a cloud environment, the critical scheduling modeling that oriented the data mining can learn from the critical scheduling method in project management. According to specific data mining workflow, combined with the control flows between tasks and the input-output relationship, the workflow was abstracted into DMTOE (Data Mining Task on Edge) network, after obtained the critical scheduling of the data mining workflow, based on multiple examples pricing model of cloud resource, it carries out the multi examples portfolio purchase.

In DMTOE network, the vertices represent events, directed edges represent tasks, the weights on the edge represent the duration of the task. The node without predecessor node was called the entrance node and the node without the successor node was called the exit node. There are multiple entrance (exit) nodes in some task graph and they can be connected to an entrance

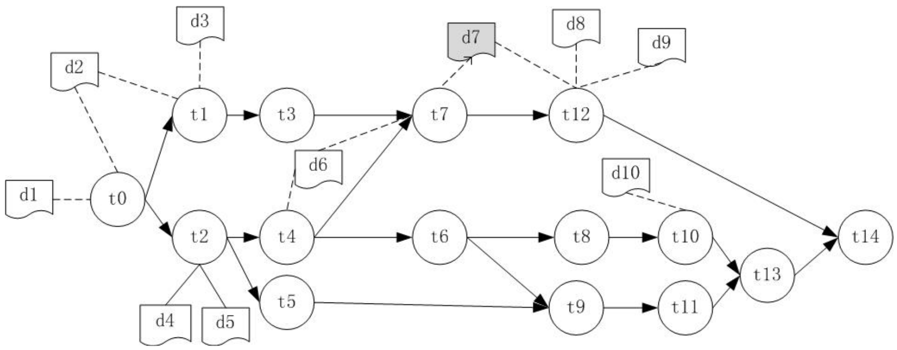

Fig. 3: Data mining workflow

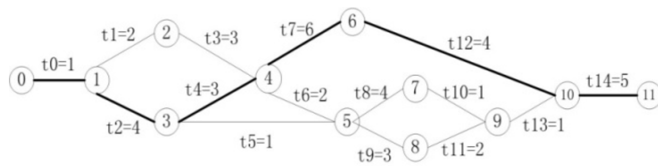

Fig. 4: Data mining critical path graph

(exit) node by the edge with the weights of 0. The directed edges described the sequence between the tasks. The node cannot be executed before receiving the message from the successor nodes and a return message from the predecessor nodes.

We assume that the data mining workflow has a total of 15 tasks, 10 data sets, 5 data centers, the dependence relationship between tasks and the call relations between data and tasks, as shown in Fig. 3. And then convert them into DMTOE network, as shown in Fig. 4.

Figure 4, node 0 is the entrance node, node, node 11 is the exit node, the weights of the edge are the execution time of a single task which got from the optimal layout scheme of cloud resources in the last chapter, according to the execution time and sequence relation of each sub task, the critical scheduling of data mining can be obtained. The critical scheduling of a data mining workflow refers to a series of tasks to decide the earliest finish time of the data mining. It is the longest scheduling in the DMTOE and has the least time to float time or the time differences. All of the tasks on the critical scheduling in DMTOE network is called the critical task and the other tasks are called non critical tasks. To solve the critical scheduling of the data mining task, we need to calculate the earliest start time, the earliest finish time, the latest start time and the end time in order. Specifically:

The earliest start time of the task (hereinafter known as ES): Refers to carry out a forward calculation through taking the expected start time as the reference points. For the task from the beginning, its earliest start time ES = 0, according to the following formula to propose a recursive calculation:

$$ES(j) = Max\{ES(i)+DU(i, j)\}, \text{ where } i \in [1, n-1], j \in [2, n]$$

DU refers to the execution time of the task, namely the weight in DMTOE network.

The earliest finish time of the task (hereinafter known as EF): refers to the earliest time to finish a task. According to the earliest start time of the data mining applications to determine the earliest finish time, namely EF = ES+DU.

The latest finish time of a task (hereinafter known as LF): refers to the latest time to complete the task within the completion time. For the final task, its latest finish time is the stipulated completion of the date mining, thus carrying out a backward calculation:

$$LF(i) = Min\{LF(j) - DU(I, j)\}, \text{ where } I \in [1, n-1], j \in [2, n]$$

The latest start time (hereinafter known as LS): refers to the latest time to start the task within the completion time. According to the latest finish time of data mining to determine the latest start time, namely LS = LF-DU.

The characteristics of the critical task is ES (i) = LS (i). And the purpose of analyzing the critical scheduling is to distinguish what is the key task in the whole data mining, strike for improving the critical task efficiency, so as to shorten the total time of data mining.

To ensure effective implementation of the tasks in the critical scheduling, buy high quality and guaranteed computing resources, the tasks in the non critical scheduling can be allowed to delay to a certain extent. And you can consider through multiple instances portfolio strategies of cloud resources to reduce the use cost. Solving the critical scheduling is the prerequisite and foundation of the multi instances portfolio purchase.

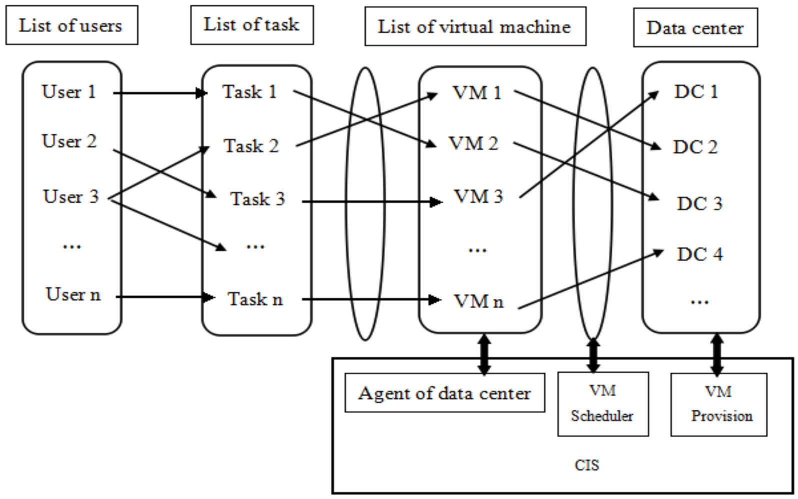

Fig. 5: Food resource scheduling model of cloud computing

RESULTS AND DISCUSSION

Design of optimization model: In this study, by using the characteristics of Map/Reduce, it can set up food resource scheduling optimization model under the cloud environment. Map/Reduce is by using the distributed parallel algorithm, to deal with a programming model for large data sets in clusters, it is also a kind of common food resource scheduling model, which is widely used, Map/Reduce mainly consists of a Map process and a Reduce process, the Map process can divide tasks into smaller sub-tasks and assign them to the working nodes, after the implementation of task, the working nodes can return the results to the main nodes, when the main nodes receive the inputting tasks. While Reduce process can have the reduction of integration in some way and then output the final result, when the master nodes receive the execution result of the working nodes. Combined with the characteristics of Map/Reduce model, this study takes the optimal time span and balance load as the goal to establish the optimal mapping relationship between tasks and virtual machine with cloud computing food resource scheduling model, which can be shown in Fig. 5.

The core idea of the optimization model is: when the user submits tasks, these tasks can be decomposed into a number of sets including a plurality of <key, value> according to different user's ID, this process is the mapping phase of the Map/Reduce model; these elements in the collection can be performed in parallel and the value of key is associated with the user's ID, so as to make response to the users who completes submitting the task scheduling, then through the food resource scheduling strategy, with the optimized targets to assign these sub-tasks to the food resource nodes of virtual machine. This process is completed by data center server agent; finally after operating the completion of the results, it can make response to users, which is the simplification stage of Map/Reduce model.

The processing of index: This study is mainly based on the fitness function of genetic algorithm to design scheduling algorithm, while scheduling algorithm in cloud environment is close related to the time span of virtual resources nodes and the balance of load, therefore, it needs to have the corresponding treatment over these two indicators, namely, time span and balance of the load balance.

Imaging, there are *m* sub-task requests on *n* food resource data centers, among them, the number of task that is assigned to each food resource is *j*, the time span that this data center can complete the task can use time cost to represent, thus, the formula of calculating time cost is as follows:

The three sub- indicators for the balance of load, namely: CPU, the unit of memory and network bandwidth is not the same, so these parameters should be normalized to one unity, so as to eliminate the differences between each sub- index in dimension.

CONCLUSION

As new type of distributed computing, cloud computing can integrate and manage various types of distributed network resources, however, during the process of unified management, the huge number, strong distribution, plus the dynamic change of food resource can have many problems. Therefore, choosing the food resource scheduling strategy to improve the organization and scheduling can have a significant meaning not only in food resource sharing and utilization efficiency, but also in saving energy and reducing operating costs. Therefore, the food resource scheduling becomes one of the key technology of cloud computing.

REFERENCES

Bezdek, J.C. and S.K. Pal, 1992. Fuzzy Models for Pattern Recognition. IEEE Press, New York, pp: 27.

Geng, Y., J. Chen and K. Pahlavan, 2013. Motion detection using RF signals for the first responder in emergency operations: A PHASER project. Proceeding of IEEE 24nd International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC, 2013). London, Britain Sep. 2013, pp: 98-107.

Li, X., Z. Lv, B. Zhang, W. Wang, S. Feng and J. Hu, 2015. WebVRGIS based city bigdata 3D visualization and analysis. Proceeding of the IEEE Pacific Visualization Symposium (PacificVis).

Mark, C., D. Niyato and T. Chen-Khong, 2011. Evolutionary optimal virtual machine placement and demand forecaster for cloud computing. Proceeding of the 2011 IEEE International Conference on Advanced Information Networking and Applications, pp: 348-355.

Tan, J., P. Dube, X. Meng and L. Zhang, 2011. Exploiting food resource usage patterns for better utilization prediction. Proceeding of the 31st International Conference on Distributed Computing Systems Workshops (ICDCSW'11), pp: 14-19.

Tseng, F.M. and G.H. Tzeng, 2002. A fuzzy seasonal ARIMA model for forecasting. Fuzzy Set. Syst., 126: 367-376.

Ye, J., 2011. Cosine similarity measures for intuitionistic fuzzy sets and their applications. Math. Comput. Model., 53: 91-97.