

Attribute Theory Model Based Task Scheduling Algorithm on Cloud

Xiaolan XIE¹, Ruikun LIU¹, Xin Hu¹, Jinsheng NI¹

¹*Institute of Information Science and Engineering, Guilin University of Technology,
Guilin, China
xie_xiao_lan@foxmail.com*

Abstract

With the arrival of the era of big data, Cloud computing has become an indispensable computational model of Internet service. As an important means of managing cloud resources, Task scheduling algorithm in Cloud directly affects the success of cloud computing. In this paper, Attribute Theory model is introduced into the task scheduling algorithm in Cloud. We establish a logical judgment standard which can indicate the resources' Qos preference of users through Attribute Theory model. And it helps users make scheduling choice in the corresponding resource with higher user satisfaction and scheduling performance.

Keywords: Cloud service, task scheduling, Attribute Theory model

1. Introduction

Cloud computing is a calculation method which can supply the virtual resource software and data of dynamic and extensible to users through the Internet^[1]. Task scheduling method in Cloud environment can reduce the performance requirements of terminal computing. The supermatic management model needs little manual intervention, which can save a large number of artificial cost . So the task scheduling in Cloud is an essential technique in the field of cloud computing. Cloud computing, if it wants to obtain the considerable development, we must pay more attention to the cloud users' Qos requirements. So it introduces a very crucial question of how to schedule tasks to uses Efficiently and rationally according to the needs of users^[2].

How to satisfy the needs of different users and map the priority of tasks and the Qos requirements reasonably is a research focus of task scheduling in Cloud computing. Document^[3] uses the DLT to design effective strategies to minimize the total processing time and to ensure that the processor is load balancing in Cloud scheduling. Document^[4] presents the work-flow QoS scheduling method based on trust relationship. This method computes the trust degree using D-S theory, and adds the trust relationship indicators in QoS factor of work-flow scheduling. This method can both meet the requirements of the Qos of work-flow scheduling and the Qos of Trust. Document^[5] presents the task classification mechanism based on SLA. And it improves task scheduling algorithm in Cloud platform. Document^[6] builds a resource scheduling model based on QoS returns in private Cloud environment through the analysis of QoS parameters and the construction of QoS profit evaluation model. These algorithms can meet the user's Qos demand in a certain extent, but most of them cannot select the appropriate resources according concrete resources Qos preference of uses. This paper presents a Attribute Theory Model based task scheduling algorithm on Cloud. The introduction of Attribute Theory Model can help multiple attribute decision making^[7]. So it can better meet the user resource Qos preference and improve the scheduling performance of Cloud.

2. Problem Formulation

There are a variety of types of resources in the Cloud environment. Different users have different preferences for these resources. Some tasks demand high performance computing resources. Some users require a lot of storage space to ensure the running of tasks. Some users are more focused on the smooth communication system. The problem of task scheduling in Cloud environment is not only solving the problem of the task scheduling deadline, but also balancing the expectation of scheduling results of users in the aspect of bandwidth, computing power and storage capacity.

The Cloud task scheduling includes 3 basic factors: Cloud resources, Cloud tasks and the scheduling strategy [8].

2.1. The Resource Model

The resources set $R=\{r_1, r_2, \dots, r_n\}$, $|R|$ is the number of resources. The resource performance of the i th resource r_i is $rCap$. $rCap$ can be further described as $rCap=\{rComp, rBW, rStor\}$. $rComp$ is the computing power of resources, rBW is the bandwidth of resources, and $rStor$ is the storage capacity of resources. $rComp \in [0,100]$, $rBW \in [0,100]$, $rStor \in [0,100]$.

Definition 1 The comprehensive performance of a resource rGP :

$$rGP_i = \sqrt{\frac{\alpha(rComp_i)^2 + \beta(rBW_i)^2 + \lambda(rStor_i)^2}{\alpha + \beta + \lambda}} \quad (1)$$

From above mentioned, α is the empirical coefficient of the computing power, β is the empirical coefficient of the bandwidth, and γ is the empirical coefficient of the storage capacity.

Definition 2 The comprehensive performance of classified resources $crGP$: It is the average value of the comprehensive performance of a certain type of resource. $crGP$ can be formulated as follows:

$$crGP = \frac{1}{n} \sum_{i=1}^n rGP_i \quad (2)$$

2.2. The Task Model

The tasks set $J=(j_1, j_2, \dots, j_m)$, $|J|$ is the number of tasks. The resource requirements of the i th task j_i is $jRRes$. $jRRes$ can be further described as $jRRes=\{jComp, jBW, jStor\}$. $jComp$ is the requirements of the computing power of tasks, jBW is the requirements of the bandwidth of tasks, and $jStor$ is the requirements of the storage capacity of tasks.

Definition 3 The expectation of the comprehensive performance of resources of tasks jGR . jGR can be formulated as follows:

$$jGR_i = \sqrt{\frac{a(jComp_i)^2 + b(jBW_i)^2 + c(jStor_i)^2}{a + b + c}} \quad (3)$$

From above mentioned, a is the empirical coefficient of the computing power, b is the empirical coefficient of the bandwidth, and c is the empirical coefficient of the storage capacity.

Definition 4 Task preference coefficient jRC : It is the quantification of the level of preference of tasks to the three types of resources. jRC is the shortest distance of jGR to $crGP$.

$$jRC_i = \min\{|jGR - crGP_i|\} \quad (4)$$

Where $crGP_i$ is the comprehensive performance of classification of resources of the three types of resources.

2.3. Attribute Theory Model

Resource performance $rCap$ is reflected by resource performance attribute $C_k(r)$. The attributes of resources are bandwidth, computing power and storage capacity. The three attributes are treated as coordinate axis. We can build a coordinate system of resource performance attributes. The 3 component vectors of the resource performance vector $rCap(r) = \{C_{rComp}(r), C_{rBW}(r), C_{rStor}(r)\}$ can form a cube. In the coordinate axis, The resource performance of each resources $rCap(r)$ can correspond the only coordinate $P_r = \{p_{rComp}(r), p_{rBW}(r), p_{rStor}(r)\}$ in the coordinate system.

Suppose that in Attribute Theory model, the resources of which the sum of the resource performance indicators that is equal to 150 can form a triangle $\triangle ABC$. By using the method of the placing of attribute coordinates, We can place the resource demand indicators of task j in attribute coordinate system which is composed by the $\triangle ABC$. Then we can get the point $P_j(p_{jComp}(j), p_{jBW}(j), p_{jStor}(j))$. P_j is the ideal demand indicators coordinate point of task j in the $\triangle ABC$. And we can calculate the distance $D(P_r, P_j)$ between the resource performance indicators point $P_r = \{p_{rComp}(r), p_{rBW}(r), p_{rStor}(r)\}$ of resource r and the ideal demand point of task j . Then $D(P_r, P_j)$ can be the comprehensive evaluation basis of task j to resource r . It is obvious that the resource r will meet the demand of task j when the value of $D(P_r, P_j)$ is low. On the contrary, the resource r will not meet the demand of task j when the value of $D(P_r, P_j)$ is high. Assume that in the $\triangle ABC$, $F(P_r, P_j)$ is the degree of fit of resource r to task j . The function relation among $D(P_r, P_j)$, $F(P_r, P_j)$ and $P_r = \{p_{rComp}(r), p_{rBW}(r), p_{rStor}(r)\}$ is as follow:

$$F(P_r, P_j) = f(D(P_r, P_j)) \quad (5)$$

Since the maximum value of three performance indicators of resources is 100, The optimal performance indicators distribution point is $P_r = (100, 100, 100)$. So $\triangle ABC$ will be gradually close to point $P_r = (100, 100, 100)$ with the increase of the performance indicators. Each ideal demand indicators points in the resource performance indicators summation section of task j will be close to $P_r = (100, 100, 100)$.

3. The Cloud Resource Decision Method based on Attribute Theory Model

3.1. The Solution of the Standard Performance Indicators Point

The different brackets of the total indicators of resource performance can generate a plurality of cross sections of the total indicators. And each section will have a standard performance indicators point. The standard performance indicators line is formed by these points. It is very difficult to express this line by common function. In order to solve this problem, we let tasks choose some desired performance indicators from the given resource performance indicators sample in the plane of the sum of the performance indicators. The selected i th sample $Rep_{rise} = (P_r = \{ap_{ricot}(r), re_{prise}(r), pr_{ehistoric}(r)\})$ will be endowed with expectations S_{pi} , and $S_{pi} \in [0, 1]$. So the standard performance indicators point in the hyper plane can be calculated by the weighted average method.

The standard performance indicators point = $\sum S_{pi} E P_{ri} / \sum S_{pi}$. Where i is the selected i th sample.

The standard performance indicator line is found out by sample study methods. We use the method of interpolation training as the sample study method.

The polynomial $y(t)$ is used as the interpolation training of the training cases.

$$y(t) = a_0 + a_1 t + a_2 t^2 + \dots + a_n t^n, \quad (6)$$

Where $aj = (a_1 j, a_2 j, \dots, a_{ij}); i = 1, 2, \dots, m; j = 0, 1, 2, \dots, n$.

Then we can get the three dimensional expression of the standard performance indicator line.

$$y(t) = y(y_1(t), y_2(t), y_3(t)) \quad , \quad (7)$$

$$\begin{aligned} \text{Where } y_1(t) &= a10 + a11 + a12t^2 + \dots + a1nt^n, \\ y_2(t) &= a20 + a21 + a22t^2 + \dots + a2nt^n, \\ y_3(t) &= a30 + a31 + a32t^2 + \dots + a3nt^n. \end{aligned}$$

The sum of the resource performance indicators G_i of the i th bracket of the section will be treated as t and fed into the standard performance indicators line $y(t)$. Then we can get the standard resource performance indicators point $I(G_i) = (y_1(), y_2(), y_3(G_i))$ in the plane of the sum of the resource performance indicators.

3.2. Determination of the Expectation

Definition 5: $f(\text{Pr}): \text{Pr} \rightarrow [0,1]$ is the expectation function. The expectation function assign numerical value to resource performance indicators. And make the numerical value situated between $[0,1]$. And the $f(\text{Pr})$ will be high with the increasing of expectation of tasks to resources.

3.2.1. Determination of the Expectation in the same Bracket: In the same bracket, The expectation of the most ideal performance indicators point I_{Pr} will be maximal.

$f(I_{Pr})=1$. So the expectation of the point will be high when it is close to I_{Pr} . In order to calculate the set F of the performance indicators of each resources, We map the three types of the performance indicators of each resources in the interval $[0,1]$ by dividing it by its maximum value. And its discourse domain U is the performance of the three types of the performance indicators of each resources. Then the set F will be the set of the optimal performance indicators. And the discourse domain $U = (k_1, k_2, k_3)$. Where $k_i \in F$ means that k_i belongs to set of the optimal performance indicators. And it is fuzzy. Define the degree of membership of C_{ki} is 1 when the value of C_{ki} is the maximum value. And the degree of membership of C_{ki} is 0 when the value of C_{ki} is 0. The degree of $k_i \in F$ can be measured by the value of C_{ki} . The set of the optimal

$$F = \left(\frac{f_1(k)}{T}, \frac{f_2(k)}{T}, \frac{f_3(k)}{T} \right)$$

performance indicators F can be shown as: ,where T is the maximum value of the performance indicators of resources. Then we can concoct the expectation by borrowing the closeness degree of the set F . The following is the definition of the fuzzy set.

Assume that $A, B, C \in F(U)$, The mapping $N: F(U) \times F(U) \rightarrow [0,1]$, Then it meets the following conditions:

- 1) $N(A, B) = N(B, A)$;
- 2) $N(A, A) = 1, N(U, \phi) = 0$;
- 3) if $A \leq B \leq C$, then $N(A, C) \leq N(A, B) \wedge N(B, C)$.

$F(U)$ is the set of fuzzy sets in the discourse domain U . The closeness degree of A and B in F is $N(A, B)$. And N is the closeness function of $F(U)$.

We can obtain the closeness degree $N(I, \text{Pr})$ of each performance indicators points Pr and the standard performance indicators point I by the fixed point I in the same plane of the sum of the performance indicators. And Pr is the expectation of this plane of the sum of the performance indicators. The expectation $f(\text{Pr}) = N(I, X)$, $X \rightarrow [0,1]$, and

$N(I, I) = 1$. It means that the optimal expectation point is the standard performance indicators point in this plane of the sum of the performance indicators. The following is the closeness degree function:

$$N(A, B) = \frac{3}{2} - \frac{1}{1 + \exp(-\sum_{i=1}^n ((A(ui) - B(ui))^2))} \quad (8)$$

The formula shows that the expectation will be small when the distance from the resource performance point to the standard performance indicators point.

Assume that (p_{i1}, p_{i2}, p_{i3}) is the resource performance of a resource R in the i th bracket, The following is the expectation $Cspi(R)$ of R:

$$Cspi(R) = \frac{3}{2} - \frac{1}{1 + \exp(-\sum_{j=1}^3 ((y_j(t_i) - p_{ij})/T^2))} \quad (9)$$

Where T is the maximum value of the performance indicators of resources.

3.2.2. Determination of the Expectation in the Overall Performance Indicators Space:

Different resources may have the same expectation in the same bracket, but their comprehensive expectation may be different. So resources will be more in line with expectation when the sum of the performance indicators is high. In this case the expectation is related to the sum of the performance indicators. So we adjust the expectation function in the same bracket by setting up the regulation coefficient r .

$f(IP_r)$ is shown as follow:

$$f(Pr) = rN(I, X) \quad (10)$$

Where $r = f(rComp + rBW + rStor)$, and the regulation coefficient r meets the following conditions:

- 1) If $rComp + rBW + rStor = 300$, then $r=1$;
- 2) r increases with the increasing of the sum of the performance indicators. The formula is as follows:

$$Ctpi(R) = (\frac{TR}{3T})^\lambda Cspi(R), \lambda > 0 \quad (11)$$

3.3. Attribute Theory Model based Task Scheduling Algorithm on Cloud

In order to gain the maximum Qos benefit in the process of scheduling, we introduce Attribute Theory Model in to scheduling algorithm. The process of Attribute Theory Model based task scheduling algorithm is as follows:

- (1) Submit the set of tasks and the set of resources to the center for task scheduling;
- (2) Save the task into the task queue by the center for task scheduling, and bracket the resources in accordance with the sum of the performance indicators.
- (3) Take out a task from the task queue, and determine the resource preference by calculating the task's resource preference coefficient jR .
- (4) Calculate the expectations of each resource in the set of resources of the definitized resource preference.
- (5) Output the appropriate resources collection to the task.

The procedure of Attribute Theory Model based task scheduling algorithm on Cloud can be shown in Figure.1:

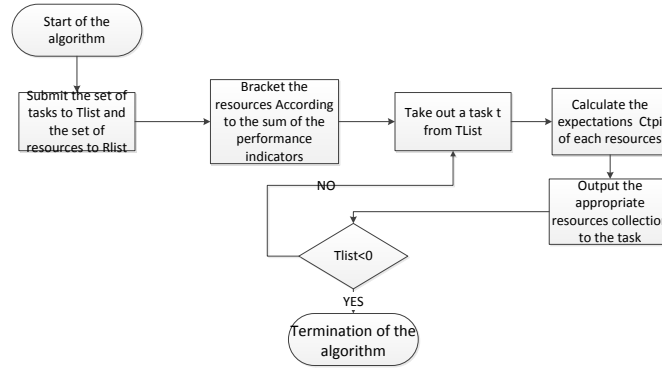


Figure 1. Attribute Theory Model based Task Scheduling Algorithm on Cloud

4. Performance Evaluation

4.1. Simulation Setup

We adopt the CloudSim^[9-11] for the simulator. It is implemented based on the programming language of JAVA. We have extended the proposed algorithm into DatacenterBroker.java for task scheduling. There are 150 virtual resource nodes. The simulation is based on desktop computer with 2.94Ghz dual core CPU, 500G hard disk and 8G memory using Linux operating systems. And we set the total number of tasks in the area between 100 to 600.

We use Task Completion Time and User satisfaction^[13] to evaluate the performance of algorithms. The smaller the Task Completion Time is, the better the performance of the algorithm is. User satisfaction is calculated by the comparison of the actual resources share and the expected resources share.

The Satisfaction of individual task Jvalue can be calculated as follows:

$$Jvalue_i = \ln\left(\frac{\mu(jComp_i)}{\sum_{k \in jORset_i} rComp_k}\right) + \frac{\omega(jBW_i)}{\sum_{k \in jORset_i} rBW_k} + \frac{\rho(jStor_i)}{\sum_{k \in jORset_i} rStor_k} \quad (12)$$

Where μ, ω, ρ is the empirical coefficient of the three types of resources. $jORset_i$ is the actual assigned set of resources. Then User satisfaction UJvalue can be calculated as follows:

$$UJvalue = \frac{\sum_{i \in uTaskSet} Jvalue_i}{n} \quad (13)$$

Where uTaskSet is the set of tasks of users, n is the number of tasks.

4.2. Comparison

The proposed algorithm ATA is implemented and compared with two other algorithms: FCFS and LBGA^[12], respectively.

As it can be seen in Figure 2, FCFS has the worst results, because FCFS does not take the demand for Qos into account. LBGA only pays attention to the number of the virtual machine's instruction execution vpMipss. It ignores the virtual machine's storage capacity vpSize and the bandwidth of the virtual machine and so on. It can not reflect the demands of the task. So it has better performance than that of FCFS. Obviously, Attribute Theory Model for ATA fully considers the Qos requirements of tasks. Therefore, it has best performance in User satisfaction.

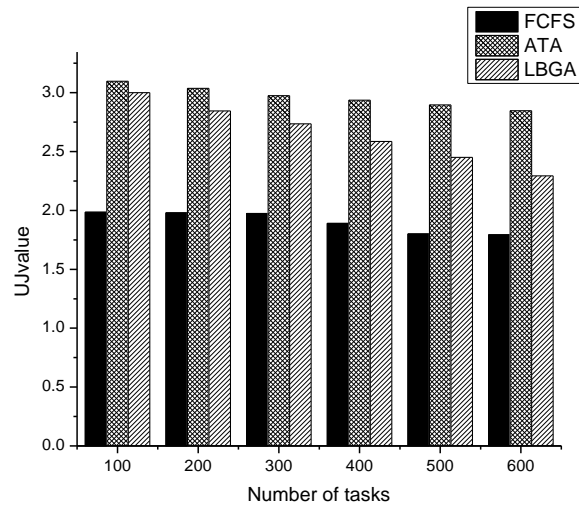


Figure 2. Comparison of the User Satisfaction

Due to the importance of task completion time in task scheduling . We also take task completion time for comparison. In Figure 3, FCFS has the worst completion time due to it allocates resources blindly. So with the increasing number of tasks, the task completion time obtained by FCFS is longer than that of LBGA. But there are some jitter in LBGA. ATA takes the resource requirements of tasks into account, it better reflects the principle of distribution according to need. So it shortens the task completion time on the whole.

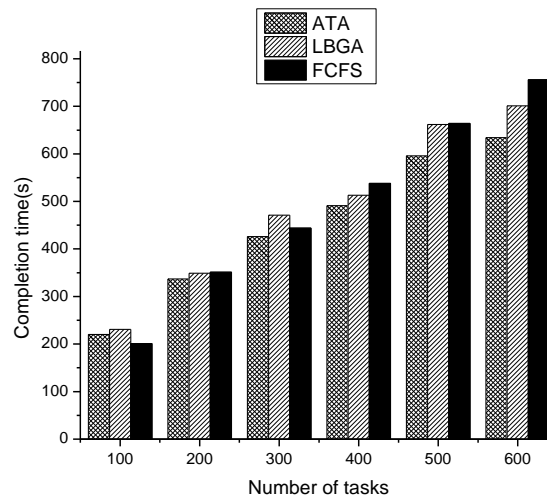


Figure 3. Comparison of the Task Completion Time

To sum up, Figure 2 and Figure 3 have shown the effectiveness of the proposed algorithm. It not only guarantees the User satisfaction in task scheduling, but also obtains a reasonable task completion time.

5. Conclusion

This paper proposes a Attribute Theory Model based Task Scheduling Algorithm on Cloud. With the concern of Qos requirements, the proposed algorithm significantly improves the User satisfaction of task scheduling on Cloud. Also, due to the advantages of Attribute Theory Model, the proposed algorithm is fast and low-complexity that obtains

the optimal results in global view. Finally, the CloudSim is extended with the proposed algorithm. In comparison of classical algorithms: FCFS and LBGA, the experimental results have shown the effectiveness of the proposed algorithm.

Acknowledgments

This paper was supported by National High Technology Research and Development Program 863 under Grant No. 2013AA12A402, National Natural Science Foundations of China under Grants No.61540054 and Natural Science Foundation of Guangxi Provincial under Grant No. 2013GXNSFAA019349.

References

- [1] FANG Wei, WEN Xuezhi, Cloud computing: Conceptions, key technologies and application[J], Journal of Nanjing University of Information Science and Technology: Natural Science Edition, 4(4):351-361 (2012)
- [2] WU Xiaonian, DENG Mengqing, A task scheduling algorithm based on Qos-driven in Cloud Computing [J], Procedia computer Science 17,pp. 1162-1169 (2013)
- [3] ZHAO Zhen, ZHANG Longchang, Research on Cloud Model with QoS Applied to Webucation[J], Microcomputer Applications, 30(1):163-166 (2014)
- [4] HU Chunhua, WU Min, LIU Guo-ping. QoS Scheduling Based on Trust Relationship in Web Service Workflow[J]. CHINESE JOURNAL OF COMPUTERS, 32(1):42-53 (2009)
- [5] ZHANG Huiyan, Research on Multi-Priority Job Scheduling Algorithm based on Qos Constrained under Cloud Computing Environment[D], LIAONING University (2013)
- [6] WEN Xiaolong, Research on Resource Scheduling Mechanism Based on QoS Benefit in Private Cloud[D], YUNNAN University (2013)
- [7] XIE Xiaolan, FENG Jiali. Decision of selecting matriculates based on attribute theory[J], JOURNAL OF GUILIN INSTITUTE OF TECHNOLOGY, 21(4) (2001)
- [8] Wu Yu-Qi , Zeng Guo-Sun, Trend Analysis for Scheduling Algorithm in Cloud Computing[J], Microelectronics & computer, 29(9),103-108 (2012)
- [9] Wang Xia-Jun, Study and Application of A Toolkit for Cloud Computing Simulation--CloudSim[J], Microcomputer Applications, 29(8),59-61 (2013)
- [10] Wang Yan-Ni ,Wu Wen-Hui , The Analysis of Simulation Process of Cloudsim3.0[J], SOFTWARE, 35(4), 63-64 (2014)
- [11] Cha Ying-Hua , Yang Jing-Li , Application of Cloudsim in Research of resource allocation[J]. Software Guide, 11(11), 58-59 (2012)
- [12] Liu Peng , Cloud computing (Second Edition), [M]. Beijing: Publishing House of electronics industry, 3-15 (2011)
- [13] Li Wen-Juan , Research on Trust-based Cloud Service Management Model and Scheduling Algorithms, Zhejiang University (2012)

Authors



Xiaolan Xie, she got her PhD in Xidian University, ShanXi, China. She is a Professor in School of information science and engineering, Guilin University of Technology. Areas of research include Cloud computing, Grid computing and Intelligent decision system. She is a committee member and deputy secretary general of Cloud computing expert committee of China communication society. She is also a member of China computer society CCF and IEEE CS.



Ruikun Liu, he received the Bachelor of Engineering degree in School of Information Engineering from ZhengZhou University, China in 2013, Currently he is staying at Guilin University of Technology for master degree of Computer science and technology, His research interest is Cloud computing.



Xin Hu, he is the guest professor of Guilin University of Technology, Deputy director and chief engineer of Xi'an Institute of Surveying and mapping .



Jinsheng Ni, he is the guest professor of Guilin University of Technology, and the Chairman of Beijing Oriental Titan Corp.

