

VM Hotspot Eliminating Method Based on Hot Degree Comprehensive Evaluating

Yan Yong-ming¹, Guo Jun¹, Mo Yu-yan¹, Zhang Bin^{1*} and Han Qing-liang²

¹College of Information Science and Engineering, Northeastern University

²Software Testing Center of Shandong Province

*zhangbin@ise.neu.edu.cn

Abstract

Evaluating the hot degree of VM (Virtual Machine) is the critical link of the VM hotspot eliminating process. Based on the traditional research, this paper creates hot degree evaluation system combining VM service availability and service quality factors. It proposes the subjective weight algorithm based on fuzzy analytic hierarchy process and the objective weight algorithm based on maximum deviation, and hot degree comprehensive evaluation method based on these two weights. A rule of VM cold-spot or hotspot evaluation based on hot degree is presented. The experiment result indicates the accuracy of the method in this paper is higher than traditional approaches to evaluate the hot degree, which can efficiently reduce improper migration, with smaller price.

Keywords: Cloud Computing; Hotspot Eliminating; VM Hot Degree; Cold-spot; Fuzzy Analytic Hierarchy Process

1. Introduction

In cloud computing environment, service providers should ensure the quality of services to meet service level agreements (SLA) signed with the users. The service always deployed in some VM, whose load will change, when the number of users increases or the access number increased dramatically [1]. This phenomenon will cause VM hotspot, which means the VM resources are insufficient to meet the needs of service performance [2]. Therefore, eliminate VM hotspot have a very important significance to protect the quality of service.

Most existing studies eliminate hotspot with methods based on VM resizing [3] or VM migration [4-6]. Although the resizing method cost small, and can finish the VM resizing without service interruption, but it is limited by the surplus resources in the PM. Therefore, in many cases we use method based on VM migration [4]. As the migration method cost greatly[7-9], how to reduce the inappropriate migration, which means how to recognize the hotspot VM and judge whether the VM needs to be moved, will be one important question facing the hotspot eliminate technology based on VM migration.

However, in the current study, hotspot judgment always set one or more thresholds, when the monitoring value or predictive value (such as resource utilization, system load) exceeds the set threshold, judge it as hotspot [2, 10, 11]. The fundamental of this kind of method is judge according to the single index, then combine them, without considering the relationship between each judge index, it is difficult to establish the relationship between different combination of index and the actual service quality, and also did not consider the different indexes have different adaptability to different services. Therefore, the resource usage, service availability, quality of service and other factors should be considered to make comprehensive

evaluation and judgment, namely judge hotspot based on VM hotspot method, then use migration to eliminate hotspot.

Aim at the deficiencies of existing methods, this article established elastic heat evaluation index system which can adjust according to the actual situation. Then, based on the evaluation index system, separately get subjective weights based on Fuzzy AHP and objective weights based on maximizing deviations, and the comprehensive evaluation hot degree judge method which combining subjective weights and objective weights, and get the judgment rule based on hot degree of VM and VM migration decision strategy. The content of this article is a comprehensive analysis of the hotspot related indicators, can reduce unnecessary hotspot mitigation which due to without considering the relationship between factors; on the other hand also addresses uncertainty problems because of expert subjective in AHP, make up the lack of single subjective weighting method.

The second section gives the basic process of the hot degree comprehensive evaluation method; third section describes the hot degree evaluation index system; the fourth section gives the calculation method of hot degree comprehensive evaluation value; the fifth section gives the hotspot judging rules and VM migration decision strategy; the sixth section proposed the corresponding experiments; finally summary the whole article and make the outlook for the next step.

2. VM Hot Degree Comprehensive Evaluation Method

Figure 1 is the basic process of VM hot degree comprehensive evaluation. In the process, we firstly analyze hot degree evaluation related influencing parameters, which can create a general hot degree evaluation indexes system. Based on the system, subjective weight and objective weight of each index of hotspot are calculated to get comprehensive evaluation index of hot degree. Because of the fuzziness of index evaluation, fuzzy analytic hierarchy process based on the triangular fuzzy number is adopted to determine the subjective weight of every affecting hot degree index. And then the multi-objective decision-making method is used to determine the target type of each index. Dimensionless processing is to get optimal relative degree of each index and its matrix. On this basis, maximizing deviation method is applied to determine the objective weight of every index. Finally synthesizing the subjective weight and objective weight, linear weighted sum method of multi-objective decision is used to obtain the pros and cons sort of each VM and its level.

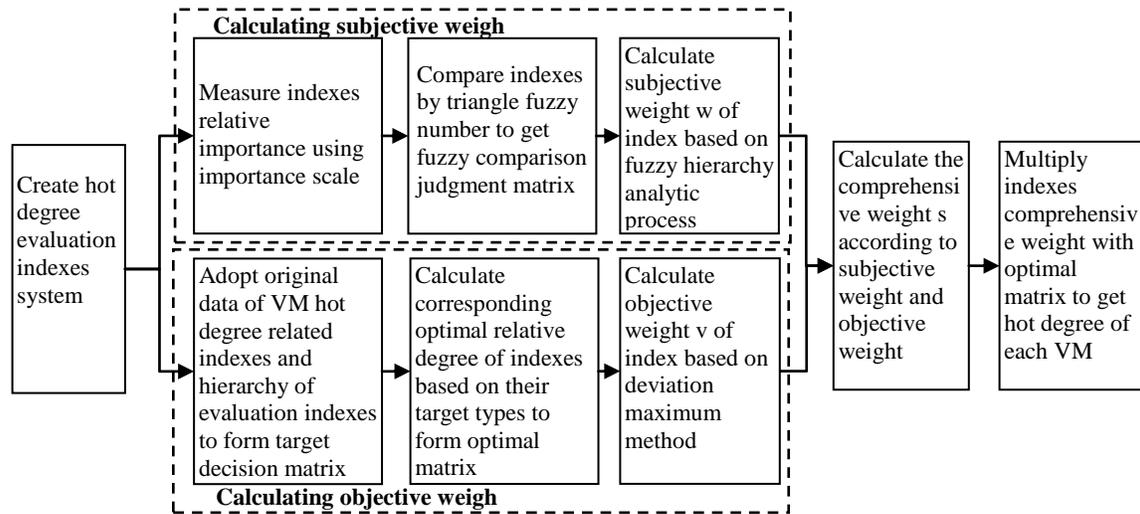


Figure 1. The Basic Process of VM Hot Degree Evaluation

3. Hot Degree Evaluation Indexes System

3.1. Service Availability Index

Service availability is the most basic parameter of service quality [4], meaning the time percentage of being able to use business functions in a given time period. Within the service time specified by SLA, Service Unavailability percentage (SUA%) is the proportion of all failures time. In this paper, Service Degradation Factor (SDF) [12] was introduced into the calculation of SUA%, being to indicate the influence of different fault time for service. And the SDF needs preliminary definition for SDF values of different failure events. Therefore, SUA% can be calculated by formula (1).

$$SUA\% = \frac{(\sum \text{Service failures time}) \times SDF}{\text{Service time specified by SLA}} \times 100\% \quad (1)$$

3.2. Service Quality Index

Performance is an important aspect of service quality, and service performance can generally be weighed through service response time, service stability, throughput, etc. Service stability can be denoted indirectly by the variability of response time, which is the deviation relative to average response time. The performance index differs from response time, and even if the response time is long, low response time variability can also help users to establish stable rhythm. Its formula is defined as follows:

$$SS = 1/q_{dev}(s) = \sqrt{n - 1 / \sum_{i=1}^n (T_i - T_\theta)^2} \quad (2)$$

T_i is response time of each calling the service, and T_θ is the response time average of calling the service.

3.3. VM Resource Utilization Index

Under the condition of ensuring the response time, service utilization is the maximum utilization that the service can be achieved, which can be said by the ratio between the used resources and the total resources in a period of time. Service utilization can show the service busy situation in a period of time, and is a key factor determining virtual machine hotspot. The utilization can determine the usage of the allocated resources to the physical node according to the resource utilization. Utilization of this virtual machine can be used as a virtual machine resource utilization threshold, and then we judge the utilization rate of resources beyond the threshold to get the resource usage. For example, if using a CPU and memory intensive applications, and more concerning about violation situation of resource utilization (beyond the initial threshold), the SLA violation ratio of the CPU and memory resources can be as evaluation index of the resource class index.

3.4. Hierarchical Structure of Hot Degree Evaluation System

The hot degree evaluation index based on cloud service quality contains resource utilization, service availability and service quality. According to SLA performance requirements of the service and completeness, independence, objectivity and other principles of indexes, the three-level hierarchical index system is established to estimate hot degree, as shown in Figure 2.

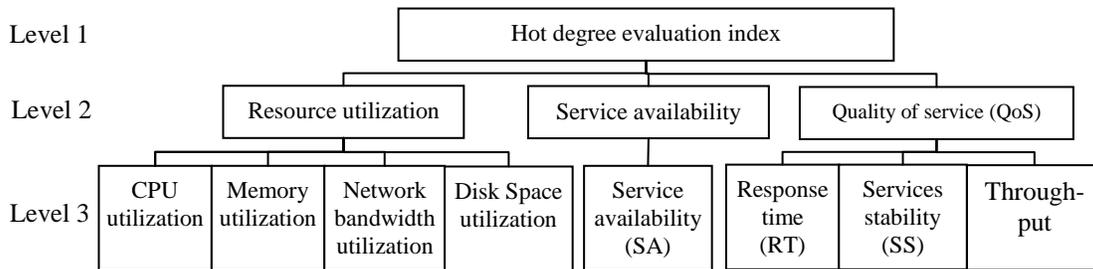


Figure 2. Hierarchical Structure of Hot Degree Evaluation System

Level 1 is called object layer, the top layer of the whole evaluation index system, which represents the final evaluation goal of the assessment strategy. And level 2 and level 3 are criterion layer and sub-criterion layer respectively. The criterion layer maps the specific requirements of the object layer to three index types, and integrates the similar parameters of the sub-criterion layer into a category of the upper layer. This index system is a universal collection of VM hot degree evaluation parameters. When concretely analyzing the hot degree of the virtual machine with some service deployed, we can specifically adjust and complement the hot degree index system.

4. Hot Degree Comprehensive Evaluation Value Calculation Method

4.1. Subjective Weight Calculating Method of Every Index based on Fuzzy Analytic Hierarchy Process

In this paper, subjective weight method of every index based on fuzzy analytic hierarchy process mainly contains three steps as the following:

(1) The Relative Importance Index of Indicators

The judgment matrix, which can be obtained by paired comparison, represents the relative importance of an upper layer element and this layer element (or the upper factors). Importance scale is a quantitative said for importance comparison results of indicators. The 1-9 scale type importance scale is applied here and its specific meaning is shown in Table 1.

Table 1. The Importance and the Meaning of Scale Type

Importance scale	signification
1	Equally important
5	Obviously important
9	Extremely important

(2) The Establishment of Expert Fuzzy Comparison Judgment Matrix

Based on the hot degree evaluation index system, triangular fuzzy numbers is introduced to form the fuzzy judgment matrix according to the uncertainty of subjective judgment experts built comparative judgment matrix. These scales order and obtain weighted value could be determined by using the theory of fuzzy number comparison size.

Assume that there are n hotspots comprehensive evaluation index of this layer related to upper layer, indices sets $A=\{a_1,a_2,\dots,a_m\}$. Triangular fuzzy number, $b_{ij}=[l_{ij},o_{ij},u_{ij}]$, is an importance fuzzy judgment of index i relative to index j experts making. l_{ij} and u_{ij} represent the fuzzy extent of judgment, the grater $u_{ij}-l_{ij}$ means the higher comparative fuzzy degree. Finally the fuzzy comparison judgment matrix B via comparison is being achieved. Similarly, the application of the method above can construct the fuzzy comparison judgment matrix of each layer relative to it's upper among index systems.

(3) The Index Weight Calculation of this Layer Relative to the Upper Layer

In the fuzzy comparison judgment matrix, the fuzzy relative weight of index i compared with other index in this layer is:

$$Q_i = \left[\frac{\sum_{j=1}^n l_{ij}}{\sum_{i=1}^n \sum_{j=1}^n u_{ij}}, \frac{\sum_{j=1}^n m_{ij}}{\sum_{i=1}^n \sum_{j=1}^n m_{ij}}, \frac{\sum_{j=1}^n u_{ij}}{\sum_{i=1}^n \sum_{j=1}^n l_{ij}} \right] \tag{3}$$

Each triangle fuzzy number in the fuzzy relative weight vector is required to clarity before sorting the current layer index. The corresponding subjective weight of $Q_i=(l_i,o_i,u_i)$ can be determined as follows [10,13]:

$$w_i = (l_i + 2m_i + u_i)/4 \tag{4}$$

4.2. Objective Weight Calculating Method of Every Index Based on Maximizing Deviations Method

(1) The Established of Target Decision Matrix

To evaluate m hot degree of n monitored spots, firstly the target decision matrix F is formed according to the measured value of monitoring spots. Meanwhile, in order to make the optimization results directly reflect the hotspot level of virtual machines, the hierarchy of hot degree evaluation index can be added into the hot degree assessment as a monitoring data

being processed. Hot degree evaluation index hierarchy can be given by experts according to the contract of service users and service providers, and the form is shown in Table 2.

Table 2. The Example of Hot Degree Evaluation Indexes' Boundaries Rating

Hot degree index	index hierarchy			
	I (cold-spot)	II	III	IV (hotspot)
CPU utilization (%)	≤10	≤35	≤70	>70
Memory utilization (%)	≤10	≤35	≤60	>60
Response time(ms)	≤10	≤25	≤50	>60

Provided with y hierarchies, the target decision matrix is established according to the assessment hierarchy thresholds and monitored data.

$$F = \begin{bmatrix} f_{11} & \cdots & f_{1y} & f_{1(y+1)} & \cdots & f_{1(y+n)} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ f_{m1} & \cdots & f_{my} & f_{m(y+1)} & \cdots & f_{m(y+n)} \end{bmatrix} \quad (5)$$

f_{ij} is the j^{th} hot degree index hierarchy threshold of the i^{th} index or the monitored value in the $(j-y)^{\text{th}}$ VM.

(2) Optimal Relative and Classification of Index Target Type

Optimal relative is the degree of relative to “optimal”, which is similar to the concept of membership degree in fuzzy mathematics, and is determined according to the target type and characteristics. The target types of single hot indexes can be divided into fixed type, range type, cost type and efficiency type. Some representative index types of hot degree evaluation can be summarized as shown in Table 3.

Table 3. The Target Type of Hot Degree Evaluation Indexes

Level 1 type	Level 2 type	Level 3 type	Target type
Hot degree	Resource utilization	CPU utilization	range type
		Memory utilization	range type
		Response time	cost type
	Service performance	Throughput	efficiency type
		Service stability	efficiency type
	Availability	Service availability	efficiency type

According to the target type of hot index, the target decision matrix converts to optimal relative matrix μ :

$$\mu = \begin{bmatrix} \mu_{11} & \cdots & \mu_{1y} & \mu_{1(y+1)} & \cdots & \mu_{1(y+n)} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mu_{m1} & \cdots & \mu_{my} & \mu_{m(y+1)} & \cdots & \mu_{m(y+n)} \end{bmatrix} \quad (6)$$

(3) The Objective Weight Calculating Method based on Maximizing Deviations Method

The index i will slight impact on the hot degree evaluation and should be given less weight corresponding if there are small differences in the hot degree i of the monitoring. Based on the optimal relative matrix, employs the method of multi-objective decision-making objective

determining weights and gets objective weight v_i of index i :

$$V_i = \frac{\sum_{j=1}^n \sum_{k=1}^n |\mu_{ij} - \mu_{jk}|}{\sum_{j=1}^m \sum_{j=1}^n \sum_{k=1}^n |\mu_{ij} - \mu_{jk}|} \quad (7)$$

Here, j and k represent different VMs respectively. $|\mu_{ij} - \mu_{jk}|$ refers to the absolute value of the optimal degree deviation between monitored spot j and k .

4.3. The Hot Degree Comprehensive Evaluation Value Calculating Method

The comprehensive weight of index i is obtained by the subjective weight w_i and objective weight v_i .

$$S_i = \frac{w_i V_i}{\sum_{i=1}^m w_i V_i} \quad (8)$$

The linear combination of the individual index evaluation value reflects the quality of comprehensive index, but there are one more indexes not qualified in the index evaluation process. An agreement is made that the comprehensive evaluation value is not qualified when any of the indicators of evaluation value obviously unqualified. For example, setting the CPU resource utilization threshold is 75%, if the resource utilization rate reached 90%, which is beyond the threshold of 20%, comprehensive evaluation value is judged to be unqualified because the CPU resource utilization index is not qualified.

The linear weighted sum method of multi-objective decision method is applied to get the hot degree evaluation index hierarchy and the value set of VM hot degree comprehensive evaluation: $E=[e_1, \dots, e_y, e_{(y+1)}, \dots, e_{(y+n)}]$, e_j is shown in the formula (9).

$$e_j = \begin{cases} \sum_{i=1}^m S_i \mu_{ij} & , \text{all indexes are qualified} \\ 0 & , \text{one or more indexes are unqualified} \end{cases} \quad (9)$$

$j=1, 2, \dots, (y+n)$. The hot degree of each VM can be sorted according to the comprehensive evaluation value.

5. Hotspot VM Judging Rule and Immigration Judging Strategy

5.1. Hot Degree Sorting and Hotspot or Cold-Spot Judging

Based on the hotspot comprehensive evaluation value, hot degrees of all monitoring point can be linear ordering. Hierarchy is also obtained in the sorting process. The specific practices are as the following: Hot degree range of each monitoring point can be obtained after uniformly ranking these hotspot comprehensive evaluation values and all levels of comprehensive evaluation values; Hotspot level of each monitoring point will be correspondingly obtained based on removing the principle; The cold/ hot judgment of the monitoring will come out in line with the level of ownership rules in advance. As a hypothesis, the set of hot degree comprehensive evaluation value E is calculated in accordance with the above method.

$$E = [e_1, \dots, e_y, e_{(y+1)}, \dots, e_{(y+n)}] = [0.8522, 0.5940, 0.2587, 0.7561, 0.4590, 0.2163]$$

The former three values are respectively the former three boundaries comprehensive evaluation values in Table 2, and the latter three are the hot degree comprehensive evaluation

values of virtual machines. The order of three VMs hot degree and the hot degree level assessment value set in advance is: Level I > VM1 > Level II > VM2 > Level III > VM3.

According to the hot degree levels in Table 2, if one belongs to Level IV it is judged to be a hotspot, so VM3 can be determined as a hotspot. Meanwhile, if no VM is in Level I, there is no cold-spot.

5.2. VM Migration Decision Strategy based on Hot Degree Comprehensive Evaluation

In order to avoid improper migration caused by a large instantaneous concurrent business, it will need to consider the time of the VM hot degree evaluation value belonging to hot level for a period of time. In a time period of n , if the VM has k time points with the hot degree evaluation value belonging to hot level, and also belonging to hot level in the next time point, we can argue that the hotspot VM needs to be migrated.

The time series prediction method based on autoregressive model is adopted to estimate the VM hot degree evaluation value in the next time point. And the values of n and k also determine that the hotspot VM migration decision strategy is positive or conservative. Given the value of n , it will produce a positive hotspot VM migration decision strategy with a smaller value of k . When $n = k = 1$, the migration decision strategy is most active, namely, when the VM is regarded as a hotspot VM, the VM will be migrated. The larger value of k is representing the VM needs being regarded as hotspot VM in more time points, which is a conservative decision strategy. A bigger value of n means a longer period of time, which is a more conservative hotspot VM migration decision strategy.

6. The Comparative Experiments and Analysis

In this section, we will compare the results of hotspot decision method based on single-threshold, multi-threshold, Fuzzy TOPSIS [5] and hot degree. The method based on single-threshold regards the VM as a hotspot when its CPU utilization is more than 60%. The method based on multi-threshold is when the CPU utilization and memory utilization are both over 60%. And the method based on Fuzzy TOPSIS algorithm calculates the workload level from CPU utilization, memory utilization, QoS metrics and so on, with the result more than 0.6. And the hot degree evaluation index hierarchy bound is shown in Table 4.

Table 4. The Boundaries Level of Hot Degree Evaluation Indexes

Hot degree index	Index bound level			
	I (cold-spot)	II	III	IV (hotspot)
CPU violation rate (%)	≤ 10	≤ 35	≤ 60	> 60
Memory violation rate (%)	≤ 10	≤ 35	≤ 60	> 60
Response time (ms)	≤ 10	≤ 25	≤ 40	> 50
Service stability	≥ 0.04	≥ 0.033	≥ 0.025	< 0.015
Service availability (%)	≥ 99	≥ 98.5	≥ 98	< 98

In the experiment, there are 2 physical machines, PM1 and PM2. We only deploy a virtual machine VM1 on PM1. And a CPU and memory intensive application is deployed on VM1. PM2 is the migration target host of VM1. LoadRunner is to produce different concurrent loads. With the given load, we respectively take some tests of the hotspot judging methods based on single-threshold, based on multi-threshold, based on Fuzzy TOPSIS, or based on hot degree, and record the results of several determining methods, as shown in Table 5.

Table 5. The Result of the Experiment Data

Hotspot judging method	VM migration times	Service availability
Single-threshold	3	95.8%
Multi-threshold	3	95.8%
Fuzzy TOPSIS	1	98.7%
Hot degree	0	100%

It can be seen that the methods based on single-threshold and multi-threshold have not considered the relationship between service quality and various factors, triggering the unnecessary migration of virtual machines every time of load peak. The latter two methods comprehensively consider resource utilization, service quality and the like, which reasonably reflects the actual operation situation of the service, and effectively avoid the improper migration, so as to ensure the high service availability. From the experimental results, the method based on hot degree is more superior to the method based on Fuzzy TOPSIS.

In addition, the method based on single-threshold and multi-threshold only need a simple comparison between monitored values and thresholds, so the time of algorithm consumption can be basically negligible. But the latter two methods need a large amount of calculation, and the computation time in the cases of different VM numbers is shown in Figure 3.

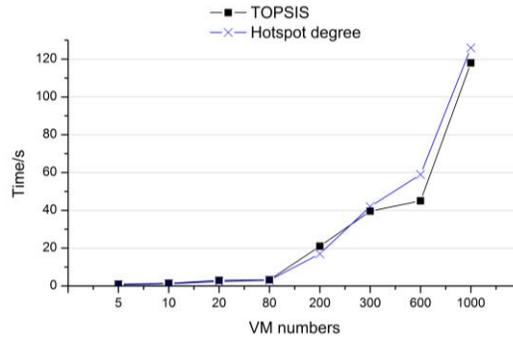


Figure 3. The Comparison of Time Consumption

We can find that the methods based on Fuzzy TOPSIS and based on hot degree can get faster calculation results with less VMs, but with more VMs, the method based on hot degree consumes a little more time, which is not obvious.

7. Conclusion

A hotspot elimination method based on virtual machine hot degree comprehensive evaluation is presented. It firstly combines the availability and quality of service deployed in virtual machine to create hot evaluation index system, as well as proposes the determination algorithm of subjective and objective weights. We then introduce the hot degree comprehensive evaluation method combining the subjective and objective weights and finally advance VM hot judgment rule and migration decision strategy. Experiment shows that based on heat hotspot judging method has better accuracy than on the single threshold and multi threshold methods, can reduce unnecessary migration in the process of hotspots elimination and verifies the rationality and validity. However, the determination method of hot virtual machine just considers the index of current system, which belongs to passive hot judgment method. So the next step will study how to use the prediction algorithm to judge hot virtual machine in advance.

Acknowledgements

This work was supported by the Fundamental Research Funds for the Central Universities(No.N120804001) and the NSFC Major Research Program(61100090, 61370155).

References

- [1] K. Appleby, S. Fakhouri, L. Fong, M. Goldszmidt, S. Krishnakumar, D. Pazel, J. Pershing and B. Rochwerger, "Oceano-SLA-based management of a computing utility. Proceedings of the IFIP/IEEE Symposium on Integrated Network Management, Seattle, American, (2001) May 14-18.
- [2] M. Mishra, A. Das, P. Kulkarni and A. Sahoo, "Dynamic resource management using virtual machine migrations", Communications Magazine, vol. 50, no. 9, (2012), pp. 34-40.
- [3] H. Nguyen Van, F. Dang Tran and J. M. Menaud, "Autonomic virtual resource management for service hosting platforms", Proceedings of the 2009 ICSE Workshop on Software Engineering Challenges of Cloud Computing, Vancouver, Canada, (2009) May 23.
- [4] T. Wood, P. Shenoy, A. Venkataramani and M. Yousif, "Sandpiper: Black-box and gray-box resource management for virtual machines", Computer Networks, vol. 53, no. 17, (2009), pp. 2923-2938.
- [5] M. Tarighi, S. A. Motamedi and S. Sharifian, "A New Model for Virtual Machine Migration in Virtualized Cluster Server Based on Fuzzy Decision Making", Journal of Telecommunications, vol. 1, no. 1, (2010), pp. 40-51.
- [6] I. Giurgiu, C. Castillo, A. Tantawi and M. Steninder, "Enabling efficient placement of virtual infrastructures in the cloud", Proceedings of the 13th International Middleware Conference, Montreal, Canada, (2012) December 3-7.
- [7] A. Verma, P. Ahuja and A. Neogi, "pMapper: Power and Migration Cost Aware Application Placement in Virtualized Systems", Proceedings of the 9th ACM/IFIP/USENIX International Conference on Middleware, Leuven, Belgium, (2008) December 1-5.
- [8] S. Akoush, R. Sohan, A. Rice, A. W. Moore and A. Hopper, "Predicting the Performance of Virtual Machine Migration, 18th Annual IEEE/ACM International Symposium on Modeling Analysis and Simulation of Computer and Telecommunication Systems, Miami Beach, American, (2010) August 17-19.
- [9] H. K. Liu, C. Z. Xu, H. Jin, J. Y. Gong and X. F. Liao, "Performance and Energy Modeling for Live Migration of Virtual Machines", Proceedings of the 20th international symposium on High performance distributed computing, San Jose, American, (2011) June 8-11.
- [10] Z. Yang, C. Li, AL. GM. Yun and C. Liu, "A new trigger strategy based on live migration of the virtual machine", Proceedings of the 2012 International Conference on Control Engineering and Communication Technology, Liaoning, China, (2012) December 7-9.
- [11] J. Xu and J. Fortes, "A multi-objective approach to virtual machine management in datacenters", Proceedings of the 8th ACM international conference on Autonomic computing, Tuebingen, German, (2011) June 14-18.
- [12] A. Strunk, "Costs of Virtual Machine Live Migration: A Survey", IEEE Eighth World Congress on Services, Honolulu, American, (2012) June 24-29.
- [13] P. Liu, Z. Yang, X. Song, Y. Zhou, H. Chen and B. Zang, "Heterogeneous Live Migration of Virtual Machines", Proceedings of the International Workshop on Virtualization Technology, Beijing, China, (2008) June 21-25.

Authors



YAN Yong-ming, he received his MS degree in software engineering from the Northeast University in 2007. Now he is a PhD candidate at the College of Information Science and Engineering, Northeastern University. His current research interests include cloud computing and service optimization.



GUO Jun, he received his PhD degree in computer science from the Northeastern University. Now, he is an associate professor of Northeastern University. His main research interests include software testing and optimizing performance on the cloud.



MO Yu-yan, he is a MS degree candidate in computer science from the Northeastern University. His current research interests is service optimization in cloud.



ZHANG Bin, he received his PhD degree in computer science from the Northeastern University. Now he is full professor of Northeastern University. His current research interests include service computing, cloud computing and web engineering.



HAN Qing-liang, he is a associate research fellow in Shandong Computing Center. His research interesting is software testing, cloud computing, information security.

