

# Research on Evaluation Method of Virtual Machine Quantity in Cold and Hot Operation Mode Based on Reliability Guarantee

Jun Guo\*, Yi Li\*\*, Chen Liu\*\*, Zixuan Zhao\*\*\*, Bin Zhang\*

\*Software College, Northeastern University, Shenyang Liaoning Province China

\*\*School of Computer Science and Engineering, Northeastern University, Shenyang Liaoning Province China

\*\*\*School of metallurgy, Northeastern University, Shenyang Liaoning Province China

guojun@mail.neu.edu.cn, Yile@outlook.com, liuchen@cse.neu.edu.cn, 1011282890@qq.com, zhangbin@mail.neu.edu.cn.

**Abstract**—Traditional service reliability evaluation involves queuing, calculating overflow, time-out, and other problems. With the development of Cloud computing and the emergence of new applications, parallel computation, redundant backup, and other complex environments shall be taken into consideration when evaluating reliability. Therefore, it's very important to carry out accurate, comprehensive, efficient reliability evaluation under the background of large-scale and complex development of Cloud computing. In view of the above problems, this paper puts forward an evaluation method of virtual machine quantity in cold and hot operation mode based on reliability guarantee, that is, to evaluate the reliability of the system by using MDD algorithm and adjust the quantity in cold and hot backup mode when the number of the virtual machine in operation mode changes under the background of meeting the reliability requirement R.

**Keywords**— Reliability evaluation; Reliability guarantee; Cold and hot operation mode; MDD algorithm;

## I. INTRODUCTION

Reliability is an essential guarantee for the normal operation of the system. A minor mistake will arouse a major failure of the software system and lead to resource damage and business interruption. To keep high reliability of the Cloud system, in terms of cloud infrastructure, redundant backup is often used to provide one or more back-up virtual machines for related virtual machines (VM). This requires that the resource management system can adjust the dynamic resources when the resource configuration of Cloud service does not match the user concurrency, to keep the SLA penalty cost low while providing services that meet user expectations constantly. There are also some deficiencies in the current study:

### A. Slow response speed to fault tolerance or high cost in the reliability guarantee process of Cloud service

Traditional redundant backup replaces the failed VM with back-up VM directly to meet the system reliability requirements, without considering the energy consumption problem. The cold back-up VM and hot back-up VM mode proposed by Remus and others[1] fail to be combined to explore the impact of the optimal allocation of the quantity of

VM in different modes on the system reliability, although it considers energy consumption on the basis of using redundant VM.

### B. Inaccurate reliability evaluation

The existing methods use webull function to reflect the failure rate while ignoring the relevance between service concurrency and service failure, as well as the impact of software aging on service failure. Moreover, the reliability calculation is more complicated in existing methods. The complexity of both time and space is high.

All these problems will result in low reliability of the current system. Therefore, this paper raises a reliability evaluation of cold and hot backup system based on multiple-valued decision diagram algorithm. When designing the evaluation interval, this paper proposes a method to evaluate the system reliability from T to T + 1 period by separating the prediction time interval, which can effectively improve the accuracy of the evaluation method. According to current service concurrency, the quantity of VM in operation mode is adjusted. Under this premise, the number of cold and hot modes should be reasonably allocated to ensure the reliability of the whole system on the basis of meeting the reliability requirement R.

## II. RELATED WORK

### A. Reliability evaluation of backup system based on MDD

In view of the reliability evaluation of the backup system, Zhai, et al.[2] raise an approximate method based on the central limit theorem to assess the reliability of cold standby system with isomorphic components; Jia, et al.[3] put forward a new kind of reliability evaluation method of multiple state decision graph based on DB-WSS considering component degradation; Rao and others[4] raise a mixed method named MSD method to conduct a reliability analysis by combining Markov forecasting method and dynamic simulating method of the system.

### B. Reliability assurance strategy

Cloud computing has a huge user group. To avoid affecting the normal use of users, the system still needs to maintain the normal operation. Hence, the reliability guarantee of the Cloud system is very important. The reliability of the whole network system can be improved by redundancy technology. Backup modes can be divided into three parts, including cold backup, warm backup, and hot backup[5]-[7]; Levitin et al.[8] suggest a new design model of the fault-tolerant system, especially for hybrid (1-out-of-n:G) redundant system whose spare components are transmitted by state dependent standby mod. A numeric algorithm is designed to consider the reliability of the system and the expected cost. It successfully formulates and solves the optimal scheduling problem of the start-up of spare units; The SFS plan proposed by Mohammed et al.[9]-[10] adopts optimal check-pointing, redundancy, and optimal selection strategies to provide fault tolerance to ensure high availability of the system and shorter system task service completion time.

### III. MODEL INTRODUCTION

#### C. Prediction of VM failure rate based on HSMM

First of all, HSMM is established to train the parameters of the model by using the historical detection data of virtual machine, and make it reflect the running characteristics of virtual machine correctly; by monitoring the running state information of VM in real-time, a time series reflecting the state change of VM is formed. According to this time series, the future running state of VM and the failure rate in the next moment can be predicted.

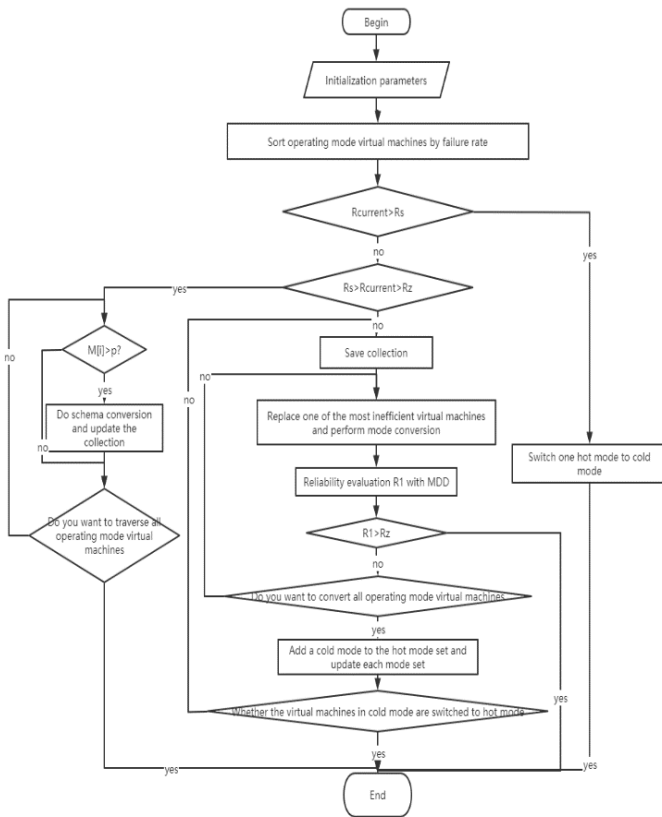


**Figure 1.** Evaluation process of VM quantity in cold and hot operation mode based on reliability guarantee

#### D. Reliability evaluation of hybrid standby system based on multiple-valued decision diagram

The running state of each VM in the Cloud service system plays a decisive role in the reliability of the whole cloud system. Based on the predicted failure rate of VM in operation mode and hot backup mode, MDD is built for operating mode virtual machine and hot backup mode virtual organization. Then MDD of the backup system is established. Finally, through the evaluation of the constructed MDD, the reliability of the backup system can be acquired.

#### E. Evaluation of the quantity of VM in hot and cold operation mode based on reliability guarantee

This paper designs an evaluation method of VM in hot and cold operation mode based on reliability guarantee. The specific flow chart is shown in Figure 1.

To ensure the reliability of cloud system within the range of user requirements, according to the performance requirements and the current service concurrency, this paper adjusts the VM quantity in operation mode and adjusts the VM quantity in cold and hot backup mode on the basis of meeting the reliability requirement  $R$ . The pseudo code of the algorithm is as follows:

##### Algorithm. Cold and hot operation mode adjustment algorithm supporting reliability guarantee

**Input:** VM failure rate matrix in current operation mode  $M$ , the reliability of current system  $R_{current}$ , VM collection in various modes of current system

**Output:** VM quantity in hot backup mode  $m$ , VM quantity in cold backup mode  $i$ , the collection of adjusted modes

**BEGIN**

1. Initialization parameters
2. Sort the failure rate  $M_o[i]$  of each OM,  $M'$  is obtained.
3. **IF**  $R_{current} > R_\delta$
4. Turn a HSM VM into CSM
5. **ELSE IF**  $R_\delta > R_{current} > R_\epsilon$
6. **FOR**  $i = 1$  to  $k$
7. **IF**  $M'[i] > p_\epsilon$
8. Turn a HSM VM into OM. Restart OM and enter CSM. Update the collection of adjusted modes
9. **END IF**
10. **END FOR**
11. **ELSE IF**  $R_{current} < R_\epsilon$
12. **FOR**  $j = m$  to  $n - k$
13.  $VM'_o \leftarrow VM_{curo}, VM'_h \leftarrow VM_{curh}, VM'_c \leftarrow VM_{curc}$  // save
14. **FOR**  $i = 1$  to  $k$
15. Add a HSM VM into collection OM  $VM_{curo} = \{VM_1, VM_2, \dots, VM_k\}$
16. Restart OM of  $M'[i]$  and enter CSM  $VM_{curc} = \{VM_{c1}, VM_{c2}, \dots, VM_{ci1}\}$
17. Add a CSM VM into collection HSM  $VM_{curh} = \{VM_{c1}, VM_{c2}, \dots, VM_{cm1}\}$
18. Evaluate the adjusted system reliability  $R1$  with MDD

```

19.      IF  $R_1 > R_e$ 
20.          Update the collection of the modes, GOTO step 29
21.      END IF
22.  END FOR
23.   $m \leftarrow m + 1$ 
24.   $VM_{curo} \leftarrow VM'_{o}, VM_{curh} \leftarrow VM'_{h}, VM_{curc} \leftarrow VM'_c$ 
25.  Add a CSM VM into collection HSM  $VM_{curh} =$ 
     $\{VM_{c1}, VM_{c2}, \dots, VM_{cm1}\}$ 
26.  Update the collection of the modes
27.  END FOR
28.  END IF
END

```

First of all, the parameters need to be initialized.  $R_g$  refers to the upper limit of reliability qualification,  $R_e$  refers to the floor level of reliability qualification;  $p_e$  refers to the limitation of failure efficiency;  $k$  refers to the quantity of VM in operation mode  $m$ ;  $M_o[k]$  refers to the failure rate matrix at the next moment of VM  $k$  in the collection of OM;  $VM_{cuvo} = \{VM_1, VM_2, L, VM_k\}$  represents the collection of the current system in operating mode;  $VM_{cuve} = \{VM_{cl}, VM_{c2}, L, VM_{cil}\}$  represents the collection of the current system in hot backup mode.

In line 2, it calculates and sorts the VM in operation mode. In lines 3 to 4, it means to shut down a VM in hot mode if the reliability of the current system goes beyond the limit. In lines 5 to 11, if the current system reliability is in the normal range, it shall consider whether the VM failure rate in a single operation mode is too high. If it is, then a VM in hot mode shall be turned into operation mode and the VM in operation mode shall be turned into the cold mode to update the collection elements of various modes. In lines 12 to 29, it refers that the reliability of current system is smaller than the lower limiting value. Line 14 is to save the current set. Line 16 is to add a VM in the hot mode to the operation mode. Line 17 is to turn the VM in operation mode with the highest failure rate into the cold mode. Line 18 is to put a VM in cold mode in the collection of the hot mode. In lines 19 to 23, system reliability is reevaluated. If it meets the requirements, the collection elements are updated and the loop ends. In lines 25 to 28, it refers to no increase in VM quantity in hot mode. Because the transfer between modes cannot meet the reliability requirements, it is necessary to add a hot mode virtual machine, restore the collection, and add a cold mode virtual machine to the hot mode to update the collection.

#### IV. EXPERIMENTAL RESULTS AND ANALYSIS

A set of default experimental parameter values are shown in Table 1. To better indicate the characteristics of various methods, the loadrunner was adopted to compress the backup system in three periods, including 0~5000 requests per second under simulated pressurization for 0~4h, 5000~10000 requests per second under simulated pressurization for 5~8h, and 1000~10000 requests per second under simulated pressurization for 8~12h. By taking the adjustment method as an independent variable, three experiments are carried out in each period. For the first time, comparison method 1 is used; for the second time, comparison method 2 is used; for the third

time, the proposed method is used. Finally, the performance of each method is compared by the average request failure rate and the average response time.

TABLE 1. EXPERIMENTAL PARAMETERS

| Sign  | Illustration                                          | value    |
|-------|-------------------------------------------------------|----------|
| T     | Total duration of one experiment                      | 12h      |
| N     | Number of VM allocated by experiment                  | 20 set   |
| C     | Maximum number of VM per server                       | 8 set    |
| $S_F$ | The average failure time of each VM in the experiment | 4h       |
| MSB   | Number of standby servers                             | variable |

Three groups of comparative experiments are designed and completed, including comparison method 1 (all operating mode), comparison method 2 (traditional backup mode), and the proposed method (the adjusting method of VM quantity in cold and hot operation mode based on reliability guarantee). The experimental results are as follows:

TABLE 2. COMPARISON OF INDEXES UNDER VARIOUS ADJUSTMENT METHODS

| Adjustment methods | The average request failure rate (%) | The average response time (s) |
|--------------------|--------------------------------------|-------------------------------|
| Proposed method    | 0.016                                | 0.108                         |
| Control method 1   | 0.031                                | 0.235                         |
| Control method 2   | 0.126                                | 0.412                         |

In table 2, the average request failure rate and the average response time under three adjusting methods are recorded. Because the proposed method adjusts VM quantity in the operation mode according to the concurrency amount at intervals and fully considers the failure of each working VM, the MDD is built to obtain the system reliability. The proposed method adjusts the operating mode virtual machine by reliability judgment so the average performance of the experimental results is the best with 0.016% of the average request failure rate and 0.108s of the average response time; the average response time of comparison method 1 is the longest for all VM resources are in operation mode; the average request failure rate and response time of comparison method 2 are lower than the proposed method. The failure rate of the VM is predicted in the proposed method. When the VM failure rate in operation mode is high, the VM is converted to the cold backup mode and the operation mode is replaced by the hot backup mode, which ensures the higher reliability of VM in operation mode. Comparison method 2 only replaces the operating system when the operation mode fails and does not evaluate the reliability of the system. Therefore, the reliability of comparison method 2 is weaker than that of the proposed method.

It can be seen from the figures that various kind of indexes are generally increasing within 12 hours by using the three methods. With the increase of running time, the system reliability decreases. The reliability of the proposed method has

little change, which proves the correctness of the method from another respect.

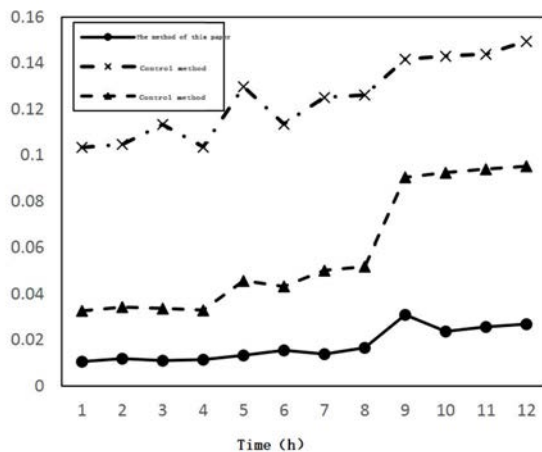


Figure 2. Average request failure rate

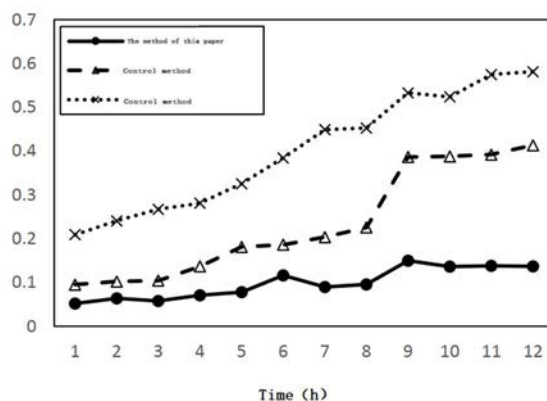


Figure 3. Average response time

## V. CONCLUSION AND PROSPECTS

In this paper, the evaluation method of VM quantity in cold and hot operation mode based on reliability guarantee is put forward to dynamically adjust the VM with operational mode failure. Through the three parallel experiments, the effectiveness of this method is verified. It not only guarantees the reliability of the system but also simplifies the calculating complexity.

While considering the reliability of a Cloud computing service system, this paper mainly takes the average request rate into account. In future work, the interruption time and other index parameters shall be considered to be more suitable for the actual environment of Cloud computing.

## REFERENCES

- [1] B. Retterath, S.S. Venkata, A.A. Chowdhury. Impact of time-varying failure rates on distribution reliability [J]. Electrical Power and Energy Systems, 2005(27): 682—688.
- [2] Zhai Q, Xing L, Peng R, et al. Multi-Valued Decision Diagram-Based Reliability Analysis of  $k$ -out-of- $n$  Cold Standby

- Systems Subject to Scheduled Backups [J]. IEEE Transactions on Reliability, 2015, 64(4): 1310—1324.
- [3] Jia H, Ding Y, Peng R, et al. Reliability Evaluation for Demand-Based Warm Standby Systems Considering Degradation Process [J]. IEEE Transactions on Reliability, 2017, 66(3): 795—805.
- [4] Meesala SrinivasaRao, Naikan V N A. A managerial tool for reliability analysis using a novel Markov System Dynamics (MSD) approach [J]. International Journal of Management Science and Engineering Management, 2009, 4(3): 230—240.
- [5] Lisnianski A. The Markov reward model for a multi-state system reliability assessment with variable demand [J]. Quality Technology & Quantitative Management, 2007, 4(2): 265—278.
- [6] Levitin G, Xing L, Dai Y. Heterogeneous 1-Out-of-N Warm Standby Systems With Dynamic Uneven Backups [J]. IEEE Transactions on Reliability, 2015, 64(4): 1325—1339.
- [7] Levitin G, Xing L, Zonouz A E, et al. Heterogeneous Warm Standby Multi-Phase Systems With Variable Mission Time [J]. IEEE Transactions on Reliability, 2016, 65(1): 381—393.
- [8] Levitin G, Xing L, Dai Y. Reliability and Mission Cost of 1-out-of-N:G Systems With State-Dependent Standby Mode Transfers [J]. IEEE Transactions on Reliability, 2015, 64(1): 454—462.
- [9] Mohammed B, Kiran M, Maiyama K M, et al. Failover strategy for fault tolerance in cloud computing environment[J]. Software: Practice and Experience, 2017, 47(9): 1243—1274.
- [10] Mohammed B, Kiran M, Awan I U, et al. Optimising Fault Tolerance in Real-Time Cloud Computing IaaS Environment [C]. IEEE International Conference on Future Internet of Things & Cloud, 2016: 363—370.
- [11] Zhao Hongwei, Shen Derong, and Tian Liling. Research on Resource Demand Forecasting and Scheduling Method in Cloud Computing Environment [J]. Mini-Micro Systems, 2016, 37(04): 659-663.



**Jun Guo**, born in 1974, Ph.D., is a associate professor at Northeastern University. His research interests include edge computing, service computing and cloud computing.



**Yi Li**, graduate student of computer technology, School of computer science and engineering, Northeastern University, From 2019 to 2022, the main research directions include edge intelligence, gesture recognition, etc.



**Chen Liu**, born in 1990, master, is an experienter at Northeastern University. Her research interests include edge computing, smart hardware and wireless sensor network.



**Zixuan Zhao**, study as an undergraduate student in major of metallurgy from Northeastern University Shenyang China. He entered Northeastern University in September 2020. His research interests include Network Communications, edge computing and so on.



**Bin Zhang**, male, born in 1964, has a doctor's degree. He is the Dean, professor and doctoral supervisor of Software School of Northeast University. He is also the chairman of the steering committee of Computer Science in Liaoning Province.