PFT-CCKP: A Proactive Cloud Services Fault Tolerance Mechanism

Jialei Liu Shangguang Wang, Ao Zhou, Fangchun Yang

State Key Laboratory of Networking and Switching Technology
Beijing University of Posts and Telecommunications
Beijing, China
{liujialei, sgwang, aozhou, fcyang}@bupt.edu.cn

Abstract—In recent years, many fault-tolerance (FT) approaches have been proposed to enhance cloud service reliability. However, exiting schemes rarely take into account coordinated problem among multiple virtual machines (VMs) which collectively complete a task, since if these VMs are uncoordinated, the execution results of a task are not correct. To solve the problem, we first exploit a proactive prediction scheme to predict the VM's status. Then, if VM's status is deteriorating, coordinated checkpoint is adopted to suspend the current task to search an optimal target host. Finally, we introduce an efficient heuristic algorithm to solve the optimal target host selection problem. The experimental results demonstrate the efficiency and effectiveness of our proposed approach.

Keywords—Cloud Service Reliability; Coordinated Checkpoint; Proactive FT; VM

I. INTRODUCTION

Cloud computing is widely applied in today's academic and industry. It enables users and businesses to use application without installation and access their own resources on any computer through Internet. Besides, cloud data center (as shown in Fig.1) mounts tens of thousands of host servers and adopts a multi-tenancy model, thus, downtime is common and has seriously affected the service level of cloud computing. So, how to enhance the cloud service reliability is a critical issue?
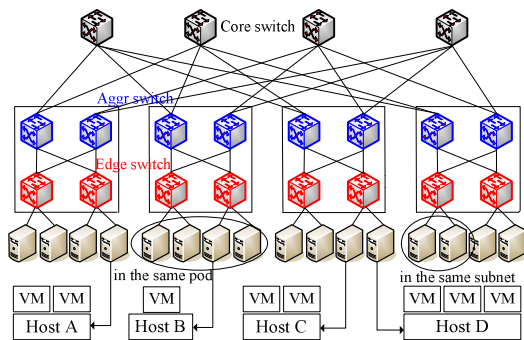

Fig. 1. Cloud data center (the switches in the top,the middle (blue),and the bottom (red) layers are the core, aggregation,and edge switches,respectively)

Then, to enhance cloud service reliability, FT which includes fault detection, backup and failure recovery is introduced. The basic idea of FT is based on exploitation of redundancy. In recent years, many FT approaches have been proposed to enhance cloud service reliability. Most of these approaches are classified into two types, one type which is widely adopted focuses on reactive scheme and relies on redundancy and checkpoint mechanism[1]. Yet, when the application behavior is very dynamic (e.g., social networks), reactive scheme can produce poor performance and may lead to low average utilization of resources. Instead of a reactive scheme for FT, the other is proactive scheme which adopts an VM prediction technique to detect a deteriorating VM executing on the cloud host. Although it reduces checkpoint frequencies as fewer unanticipated failures are encountered and is complementary to the reactive FT, it just considers a whole task to execute on a VM and does not consider a virtual cluster to collectively execute distributed applications (e.g., client-server systems and transaction processing, etc.). How to deal with this situation?

To the best of our knowledge, the checkpoint approach is widely adopted FT mechanisms. It periodically saves the execution status of a running task as a checkpoint image file, in other words, the checkpoint image only is an execution status of a VM at one point. When a virtual cluster jointly deal with a task, if one VM of the virtual cluster is fault, to guarantee status consistency of the virtual cluster, the checkpoint image cannot be exploited. So, in order to avoid complete restarts of the virtual cluster and to decrease storage and network resource consumption and recovery time because of only one failure, we adopt another typical fault tolerant technique called coordinated checkpoint. And, the global state obtained from a coordinated checkpoint is coherent, allowing the system to recover from the last full completed checkpoint in case of any failure.

In this paper, we assume that cloud data center supports the coordinated checkpoint mechanism, and introduce a proactive FT coordinated checkpoint approach, called PFT-CCKP. Firstly, a virtual cluster collectively handle a task. Secondly, when one VM of the virtual cluster is deteriorating, the virtual cluster is checkpointed and suspends the current task to search an optimal target host. Finally, the optimal target host selection problem is modeled as an optimization problem.

II. METHODOLOGY

In our proposed approach, the VM prediction and coordinated checkpoint are exploited to guarantee cloud service reliability. Then, the objective of our approach is to predict a deteriorating VM, and then to search an optimal target host for the checkpoint image of the deteriorating VM. The system architecture of our proposed approach (as shown in Fig.2) includes the following three modules.

1) VM predictor. The module adopts a adapted version of Statistical Metric model (SMM) to predict a deteriorating VM executing on a cloud host.

2) Coordinated checkpoint. When there are a deteriorating VM in a virtual cluster, the module automatically checkpoints the virtual cluster to ensure global consistency.

3) Optimal target host selection. When the module selects the optimal target host for the deteriorating VM, it is to be modeled as an optimization problem which considers three constraints including CPU power, memory capacity and disk capacity.
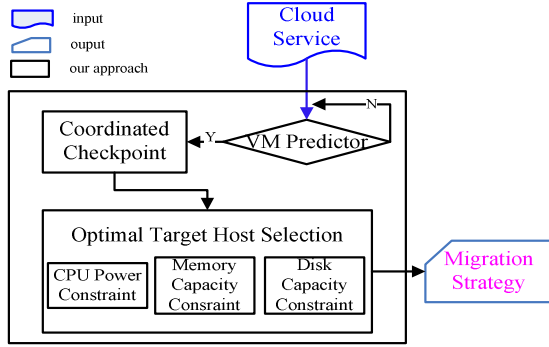

Fig. 2. PFT-CCKP architecture

In the following content, we will introduce the detail of VM predictor and the optimal target host selection in section A to B.

A. VM predictor

It is worth to mention that, The SMM which relies on Markov assumption adopts the idea of the Statistical Language Model (LM) to predict the behaviors of VMs and introduce acceptable management decision for adapting VMs.

The SMM predicts the next value for the defined criterion following the equivalent pattern based on the probability p presented by the following:

$$p(s_1,...,s_n) = \prod_{i=1}^{n} p(s_i \mid s_1^{i-1}) \approx \prod_{i=1}^{n} p(s_i \mid s_{i-n+1}^{i-1}) = \frac{c(s_{i-n+1}^{i})}{c(s_{i-n+1}^{i-1})} \quad (1)$$

where s_i^j denotes samples $s_1,...,s_n$ ($i.e.\ p(s_5 \mid s_2^4) \leftrightarrow p(s_5 \mid s_2 s_3 s_4)$), $c(x)$ denotes the frequency/count or the number of times the sequence x occurs in the metric, n is the sequence's length.

To apply the SMM model in cloud data center, in practice there are three criteria directly affecting the data-center's workload: Memory utilization U_{mem}, CPU utilization U_{CPU}, and network utilization U_{net}.

B. Optimal target host selection

In this section, The optimization objective of the optimal target host selection problem is to minimize the total distance between VM migrated and other VMs collectively completing a task in the optimization period while satisfying the resource requirements. Hence, the minimum total distance (MTDT) migrated can be modeled as the following optimization formulation:

$$MTDT = min \sum_k y_{ik} \quad (2)$$

s.t.

$$\sum_{j=1}^{n} x_{ij} = 1, \ x_{ij} = 0 \ or \ 1, \ i,k = 1,2,...,m, \ j = 1,2,...n \quad (3)$$

$$\sum_{i=1}^{m} r_i^{mem} x_{ij} < c_j^{mem} \cap \sum_{i=1}^{m} r_i^{cpu} x_{ij} < c_j^{cpu} \cap \sum_{i=1}^{m} r_i^{disk} x_{ij} < c_j^{disk} \quad (4)$$

$$\sum_j Flow_{ij} - \sum_l Flow_{li} = \begin{cases} 1 & if\ node_i\ is\ the\ node\ including\ deteriorating\ VM \\ 0 & otherwise \\ -1 & if\ node_i\ is\ the\ target\ node\ of\ deteriorating\ VM \end{cases} \quad (5)$$

where n is the number of hosts in the cloud data center and m is the number of VMs in a virtual cluster; y_{ik} is the distance between the *i-th* VM migrated to a new location and the other VM which are in the same virtual cluster with the *i-th* VM; *Eq.*3 shows that a VM can only be placed on one host, such that $x_{ij}=1$ if the *i-th* VM is run on the *j-th* host, and $x_{ij}=0$ otherwise; *Eq.*4 shows that the sum of the resource requirements for VMs must be less than the host's resource capacity; r_i^{disk}, r_i^{mem} and r_i^{cpu} are the maximum disk, memory, and CPU requirements of the *i-th* VM in an optimization period, respectively; and c_j^{disk}, c_j^{mem} and c_j^{cpu} are the disk, memory, and CPU capacity of the *j-th* host, respectively; besides, if one VM_n in the $node_n$ and the deteriorating VM_i are both in the *p-th* virtual cluster, we can assign a value p to a variable $TagNode_n$ in the VM_i and VM_n.

III. PERFORMANCE EVALUATION AND CONCLUSION

In the paper, we are promoting PFT-CCKP exploiting coordinated checkpoint which can anticipate a deteriorating VM on a cloud host, and then automatically migrate checkpoint image from source host to the optimal target host. To further study the performance of PFT-CCKP, we compare our approach with MBFD, MAD, IQR, LR, and THR in terms of network resources usage and average lost time. Experimental results based on FTCloudSim [2] show that the above metrics of our approach are better. In the future, we will apply our proposed approach to reactive FT using full coordinated checkpoint mechanism, it can not only reduce checkpoint frequencies as fewer unanticipated failures are encountered, but also reduce network and storage resources consumption while still guaranteeing cloud service reliability.

ACKNOWLEDGMENT

REFERENCES

[1] A. Zhou, S. Wang, Z. Zheng, C. Hsu, M. Lyu, and F. Yang, "On Cloud Service Reliability Enhancement with Optimal Resource Usage", Cloud Computing, IEEE Transactions on, 99, vol. PP, pp. 1-1, 2014.

[2] A. Zhou, S. Wang, Q. Sun, H. Zou, and F. Yang, "FTCloudSim: a simulation tool for cloud service reliability enhancement mechanisms", Proc.ACM/IFIP/USENIX International Middle-ware Conference, Demo &Poster Track, Dec.2013.