5X-9

An Agent-based Approach for Mining Performance Data Streams from Virtualized Environment

Shuhei Nakahashi† Khamisi Kalegele‡ Hideyuki Takahashi†,‡ Kazuto Sasai†,‡ Gen Kitagata†,‡ Tetsuo Kinoshita†,‡

† Graduate School of Information Sciences, Tohoku University

‡ Research Institute of Electrical Communication, Tohoku University

1. Introduction

Virtualization technology is becoming popular because it facilitates faster deployment. However there are still some challenging problems such as effective virtual machine (VM) resources utilization. When a VM has more resources than necessary, resources are wasted. When a VM has fewer resources than necessary, not only the performance goes down but also more energy is wasted. In order to use resources optimally, VM resources must be accurately predicted before VMs are instantiated. In this research, we present VM resource prediction system based on an multi-agent based approach. The approach uses two agents. A Streaming-agent (S-Ag) for measuring performance data and a Recommendation-agent (R-Ag) for recommending when to optimize VM resources. R-Ag uses various mining techniques (e.g. moving average) to suggest changes (*downgrade* or *upgrade*) in VM resources. For experimentation and in order to evaluate our approach, we developed a testbed system using a repository based multi-agent platform and the VirtualBox. Preliminary results show that it is possible to stream performance data from VMs and apply stream mining techniques to optimize resources utilization.

2. Problem and related work

Resource estimation for new VM provisioning is a challenging problem. Currently, in virtualized environments, users decide the amount of resources to be allocated on VMs. However, it is difficult for users to predict optimal amounts of VM resources because the optimal amounts vary depending on many factors such as provided services and workloads. Another problem is prediction of dynamic resource demand during operation. In this paper, we are focusing on prediction of dynamic resources. Although resource allocation problem has been well studied in the past, it is still an active research area in virtualized environments.

In [1], they allocate resources dynamically based on application demands and support green computing by optimizing the number of servers in use.

In [2], they propose a technique for dynamic consolidation of VMs using auto-adjustments of threshold values.

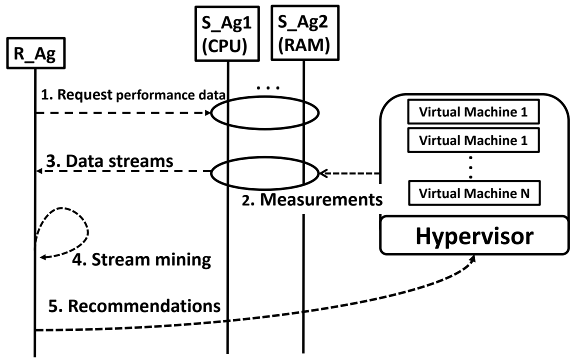

Figure 1: Agents organization

In [3], resource prediction for cloud computing has been done using double exponential smoothing. They use not only the current state of VM resources but also history records during the prediction.

Different from other existing works, in our approach we are streaming performance data from the VMs and apply stream mining techniques to optimize VM resource utilization.

3. The proposed approach

In optimizing VM resources, prediction of resources must be done within a minimum amount of time so that the predictions become relevant. Also, huge amount of performance data from the VMs must be analyzed. We propose a multi-agent based approach to optimize VM resources through prediction of dynamic resource demand. Multiple agents can cooperate among them to quicken the prediction process. We propose a design of two agents. One is Streaming-agent (S-Ag) and the other one is Recommendation-agent (R-Ag). Figure 1 shows organization of the two agents. S-Ags are knowledge based agents which measure performance data from VMs and produce data streams with specified heartbeat rate. R-Ag requests data streams from S-Ags and conducts stream mining using various mining techniques. Our case study technique is a Simple Moving Average (SMA) defined as

$$SMA = \frac{1}{w}\sum_{i=1}^{w} n_i \ ,$$

where w is window size and n is measurement data.

Other techniques which are underdevelopment include an auto regressive moving average and the Hoeffding tree algorithm based.

4. Experiment

In order to investigate the feasibility of our approach, we developed a testbed system using DASH multi-agent platform, VirtualBox and Java programming language. The VirtualBox was installed on Red Hat Enterprise Linux 6.3. The Agents were installed in a separate management machine.

For our preliminary experimentation, two VMs (VM1 and VM2) with different resources were provisioned. VM1 is running and VM2 is off. The procedure of the experiment was as follows.

(1)We put workloads on VM1 by using stress test.
(2)S-Ag gets performance data (e.g. CPU measurements) from VM1 through VirtualBox API. R-Ag requests the data from S-Ag every second.
(3)R-Ag calculates SMA with 10 past data of CPU usage. When SMA of CPU usage goes over the threshold of 65%, R-Ag sends a request to VirtualBox to switch to VM2.
(4)S-Ag and R-Ag send CPU usage data and SMA CPU usage data to Graph-agent (Graph-Ag). Graph-Ag creates a graph for those metrics.

The results of our experimentation were as follows. Figure 2 shows CPU measurement stream and SMA values. SMA values went over the predefined threshold of 65% at around 180 seconds and after a few seconds CPU usage dropped to almost 0%. At this point, R_Ag recommended a resource upgrade.

As shown in Figure 3 R-Ag sent a request to VirtualBox to switch VMs. R-Ag also sent a request to S-Ag so that S-Ag starts to stream performance data from VM2. VM1 was turned off and VM2 was turned on. S-Ag started to get CPU measurement stream from VM2.

5. Conclusion and future work

In this paper, we presented the VM resources prediction system based on a multi-agent approach and stream mining techniques. We prepared the testbed system with two virtual machines (VMs) and conducted the preliminary experiment. In the experiment, when a VM was overloaded, our proposed recommendation agent (R-Ag) recommended a resources *upgrade*. The testbed was able to switch the VM into a higher resource VM.

For our future work, we plan to investigate normal operation behavior of VMs to decide the value of measurement thresholds. We are also developing other methods based on techniques such as auto regressive moving average models and the Hoeffding tree algorithm.

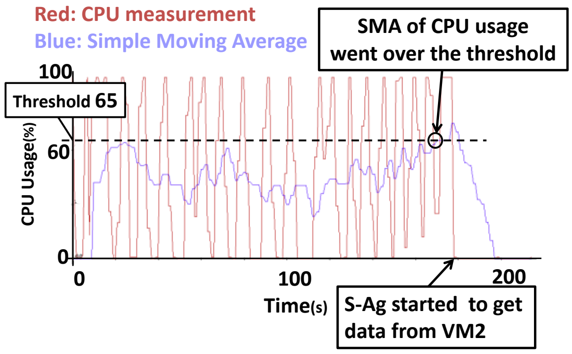

Figure 2: CPU measurement stream and SMA values

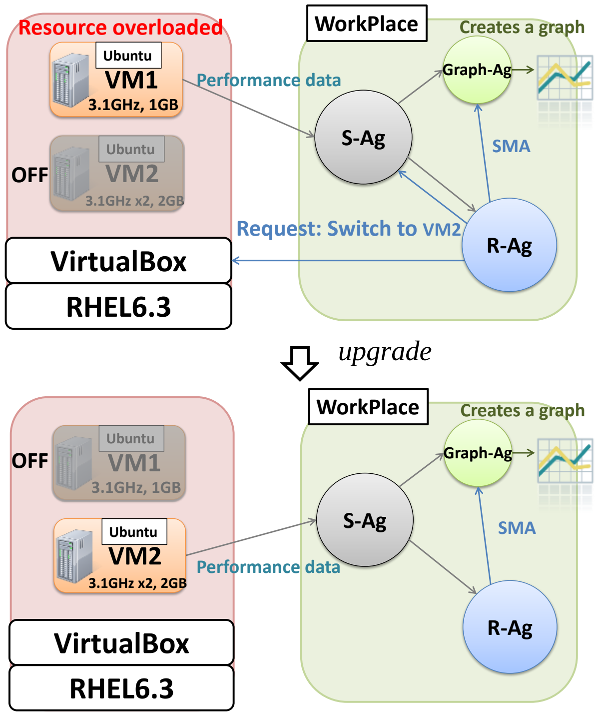

Figure 3: Resource *upgrade* experiment

References

[1] Z. Xiao, W. Song, and Q. Chen, "Dynamic Resource Allocation Using Virtual Machines for Cloud Computing Environment," *IEEE Transaction on Parallel and Distributed system*, VOL. 24, NO. 6, JUNE 2013.
[2] A. Beloglazov, and R. Buyya, "Adaptive Threshold-Based Approach for Energy-Efficient Consolidation of Virtual Machines in Cloud Data Centers," *Proceedings of the 8th International Workshop on Middleware for Grids, Clouds and e-Science*, 2010.
[3] J. Huang, C. Li, and J. Yu, "Resource prediction based on double exponential smoothing in cloud computing," *in Consumer Electronics, Communications and Networks (CECNet), 2nd International Conference on, pp. 2056-2060, 2012.*