International Conference on Computational Intelligence: Modeling Techniques and Applications (CIMTA) 2013

Agent based Resource Monitoring system in IaaS Cloud Environment

A.Meera\textsuperscript{a*}, S.Swamynathan\textsuperscript{b}

\textsuperscript{a}Department of Information Technology, Tagore Engineering College, Anna University, Chennai, 600127, India.
\textsuperscript{b}Department of Information Science and Technology, Anna University, Chennai, 600025, India

Abstract

In cloud computing environment, Infrastructure as a Service (IaaS) takes the lowest tier in the cloud pyramid where most control and management is needed. IaaS clouds offer IT infrastructure resources for computing, storage and networking to cloud users. In a real cloud data center, there are physical servers with a large number of virtual machines. These virtual machines are hosted with many heterogeneous applications. In order to optimize the utilization of computing resources and also saving energy consumption of cloud data centers, the applications running on the virtual machines will be migrated either to the same server or to another physical or virtual server. Identifying when it is best to migrate an application in a virtual machine has a direct impact on resource optimization. Performance optimization can be best achieved by an efficiently monitoring the utilization of computing resources. So, we need a comprehensive intelligent monitoring agent to analyze the performances of virtual machines. In this paper, we propose an agent based resource monitoring system that depicts the CPU and memory utilization. The monitoring agent collects the virtual machine resource usages and displays in a dashboard. Dashboard displays the key performance metrics such as CPU and memory utilization. The statistical report of dashboard provides information to cloud administrator for resource optimization.

© 2013 The Authors. Published by Elsevier Ltd.
Selection and peer-review under responsibility of the University of Kalyani, Department of Computer Science & Engineering

Keyword: cloud monitoring; resource monitoring; virtual machines; IaaS; virtualization.

Meera. Tel.: 9486203100
E-mail address: meeravishnu@gmail.com.
1. Introduction

1.1. Overview of Cloud Computing

"Cloud Computing is a model for enabling ubiquitous, convenient, on demand network access to a shared pool of configurable computing resources that can be rapidly provisioned and released with minimal management effort or service provider interaction[1]. The basic service models of cloud computing are Infrastructure as a Service (IaaS), Platform as a Service (PaaS), Software as a Service (SaaS).

Due to vast requirement of on-demand services, anything can be outsourced as a service which is commonly referred to as Anything as a Service (XaaS) [2]. Apart from basic cloud services there are other services that are offered from the cloud. They are Communication as a Service (CaaS) that outsource enterprise communication solutions to customers. The CaaS Vendors are responsible for management of hardware and software voice over IP, instant messaging services, Video Conferencing, Soft Phones, and Multimedia Conferencing. Database as a Service (DaaS) provides remotely hosted database service to users need sharing of database and make to function as if the database were local. Integration as a Service allows to access integration software functionality on a pay per use model. Governance as a Service allows the ability to manage the different cloud services.

Monitoring as a Service (MaaS) provides monitoring for security such as external threats, vulnerability detection, monitoring for trouble shooting, monitoring for Service Level Agreements (SLA) compliance and Quality of service (QoS).

The various deployment models of cloud are public cloud, private cloud, community cloud and hybrid cloud. Public cloud resources are made available to anyone over the internet. Private cloud resources are managed by a single organization. Community cloud resources are managed and shared by organizations of same community. Hybrid cloud is combination of two or more public, private or community cloud.

1.2. Need for Cloud Monitoring System

Cloud Monitoring can be referred to as monitoring the performances of both physical and virtual infrastructure. The resources that are utilized by the physical and virtual infrastructures and the applications running on them must be measured efficiently. Cloud Monitoring can be focused from different perspectives such as security monitoring to achieve confidentiality, integrity and availability of data (CIA). From technology metric perspective, resource management plays the major role.

In IaaS cloud service, a customer can rent computing resources such as processor, storage and network. A cloud data center is designed to have several racks of servers. These servers are virtualized to provide logical execution environments called virtual machines to cloud users. Cloud users run their applications on these virtual machine instances. They do not have any direct control over the physical infrastructure. This also creates a fear among them to move their applications and data to cloud. However, with proper security measures, the end user can be provided with necessary trust and confident for using the cloud infrastructure. A performance monitoring system that measures and reports the status of virtual infrastructure provides a solution to this issue.

The major problem with the traditional monitoring system is that, they provide only domain specific information. For example, a networking tool monitors only the network packets. Therefore, these performance tools report resource utilization statistics and good view of individual components. But they fail to give a good view of entire cloud infrastructure. This leads to difficulty in analyzing what is actually needed for decision making. No end user is looking for metrics. Instead, the end users look for intelligent reports to actually point to what, where and when action is needed. So, it is essential to have a monitoring system that correlates all the components and give a consolidated report about the resource availability and consumption of resources in a cloud infrastructure so that, customers believe that they pay only for what they are getting.

Another challenge a cloud consumer faces is the business responsibility [3]. Most of the cloud provider’s service level agreement says that, a customer can receive a service credit in paying the bill if their service drops below certain threshold value. But service level agreement still lacks in fulfilling many attributes related to customer’s requirement. Still in many cases, the data or business loss to the customer is not properly compensated.

On analyzing the issues, it is understood clearly that, from a technology metric perspective it should show the usage trends so that the cloud administrator can predict the tolerance capacity of virtual machines, key performance
metrics such as CPU, memory and network bandwidth utilization, metrics for troubleshooting purposes. From business perspective, a monitoring system should monitor compliance with service level agreement, track utilization with cost of service, assist in predicting the seasonal growth and maintaining maximum uptime for applications. So, a monitoring agent is essential to provide the healthy information of cloud infrastructure.

Software agent is an autonomous unit, which is capable of performing specified tasks on their own. They can achieve by being defined or learning knowledge from the environment and adapt to changes in the environment by updating new knowledge. Software agents based systems are generally classified into general purpose agents, mobile agents, learning agents, goal based agents and utility based agents. From this, one can understand that a software agent based monitoring system may provide better solution compared to other system.

The remainder of the paper is organized as follows. Section II defines the related work done by various researchers related to monitoring system. An agent based resource monitoring system architecture is defined in section III. Implementation and results are discussed in section IV and finally concluded in section V.

2. Related Work

There are some recent and related works done on performance analysis and monitoring. A survey on cloud monitoring is proposed by Giuseppe Aceto et al. They have discussed the motivations for cloud monitoring, current platforms and services for cloud monitoring [4]. Ling Liu et al. have proposed Windows based state monitoring (WISE) framework for efficiently managing cloud applications. They have proposed a system that generates alerts only when state violation is continuous within a time window to reduce the communication cost [5]. A self adaptive hierarchical monitoring mechanism for cloud is proposed by Gregory Katsaros et al. The monitoring system facilitates on the fly self configuration in terms of time intervals and monitoring parameters for film post application [6].

With the rise of load balancing techniques in cloud resources, migration of virtual machines plays a major role. Few papers are discussed based on migration of virtual machines and resource allocation mechanism. A mechanism for addition and deletion of virtual machine instance based on the load of the instance is proposed by Astuo Inomata, et al. in [7]. The system compares the current load information with the preconfigured threshold values. However, all the above mentioned work focused on dynamic adaption of Infrastructure. But these systems can be further enhanced by estimating the future demands in cloud infrastructure based the monitoring parameters. On the other hand, throughput can be improved by monitoring the shared resources across different platforms. Jaideep Moses et al. [8] have proposed a system that migrates the virtual machines that are severely affected by shared resource contention. They have focused their work in throughput optimization in cloud computing data centers. However, their work does not formulate the cost of migration. Performance overhead such as downtime, total migration time are analyzed by Keiiang Ye, Xiahong Jiang Dawei Huang Jianhai Chen, Bei Wang. Live migration is done by resource reservation methodology [9]. But their work does not depict any policy in reserving resources in a target machine for migration. Placement of virtual machines across multiple clouds using brokering architecture is explained by Johan Tordsson et al. [10].

There are also some studies conducted by researchers on agent based works. The work of Mohit Dhingra et al. introduces a distributed resource monitoring framework for IaaS cloud [11]. The virtual machine agent resides in each host reports all VM related metrics such as CPU, memory and I/O bandwidth utilization. An adaptive resource allocation using a monitoring agent is simulated [12]. Their allocation policy considers the geographical distance between cloud consumer and data center. The monitoring agent reports the availability of computing resources in a data center. End user requests are allocated to a available data center that is near to the customer’s site to reduce the network delay. Qi Liu Georgios et al. [13] presented a work on autonomic management of virtualized data centers using agents. They proposed a worker agent that monitors the cloud environment and takes its own decision when required.

Apart from monitoring the computing resources, these monitoring services can also be extended to monitor vulnerabilities. Inter-virtual machine traffic analysis for intrusion detection is presented using an autonomous agent [14]. Their work introduces a design that makes analysis on specific parts of IP packets for detecting vulnerabilities. Cellular-cloud integration framework to support real-time monitoring and management of traffic on the road is proposed by Ayalew Belay Habti [15]. Their work shows the importance of capturing live data for testing using
cloud computing. Monitoring of IaaS and SaaS layer using open source monitoring tools are discussed using a lightweight framework in public clouds [16]. So, on analyzing various research works, it is understood that monitoring is mandatory for measuring performances in cloud environment.

3. System Architecture

The architecture of agent based resource monitoring system for IaaS cloud is shown in the Fig. 1. Cloud is indeed a distributed with hybrid environment. Cloud data center is a centralized repository where computing resources such as servers, storage and network are operated and managed. Assume for a given process, virtual machines are created from physical machines by the support of virtualization. The architecture discusses only the utilization of resources with respect to various virtual machines. Generally agents are autonomous in nature. These agents are responsible for doing the assigned task. However, this paper doesn’t discuss much about the autonomous agents and will be explored with respect to static, dynamic and mobile agents in future work.

![Fig. 1. Architecture of agent based monitoring system](image)

IaaS cloud is designed to have infrastructure resources where virtual machines can be created. To enhance the monitoring services, each virtual machine is supported with a Virtual machine Resource Monitoring agent (VmRM agent). This VmRM agent collects the CPU and memory utilization of individual virtual machine hosted with different types of applications. It sends the resource usage statistics to the agent based resource monitoring system. Agent based resource monitoring system has two components. The resource usage collector component accumulates
the health information of each VM and sends that to the resource usage reporter. The resource usage reporter reports the virtual machines status information to the cloud administrator and also displayed in the dashboard. The cloud administrator has performance analyzer module which analyzes the statistical report in order to measure whether the performance is up to the expected service. The monitoring system accepts the resource status request from the cloud user and sends the status as response.

4. Implementation and Results

In order to illustrate the performance of the agent based resource monitoring system, a system with hardware configuration of CPU with 4 x 2.44 GHz, Memory (RAM) with 3.41 GB and local disk with 82 GB, installed with Linux 2.6.1.8-238.el5Xen (X86_64) was considered. Our agent is implemented using java programming. In future, we are planning to use mobile agents for resource monitoring for proper resource utilization. Eucalyptus was used to create virtual machines with a centos image customized with java and tomcat server. On top of each VM, CPU intensive application and memory intensive application such factorial calculation and finding paths in a graph were considered to test the performances of virtual machines. The CPU and memory utilization factors were discussed in table 1 using EC2 image of VM type small (1 CPU, 128 MB RAM and 3GB hard disk).

| Number of Virtual Machines and Resource Utilization % | Without Virtual Machine | With 1 Virtual Machine | With 2 Virtual Machines |
|-------------------------------------------------------|------------------------|------------------------|-------------------------|
|                                                       | CPU Idle % | Memory Usage % | CPU Idle % | Memory Usage % | CPU Idle % | Memory Usage % |
| CPU Intensive application                             | 95.5       | 60.23       | 89.42      | 74.96          | 81.42      | 96.78         |
| Memory Intensive application                          | 96.47      | 62.34       | 88.46      | 75.32          | 80.33      | 96.66         |

The results of table 1 clearly describe that, CPU utilization and memory usage increases with the increase in number of virtual machines. Since virtual machines require more memory to create, the linux cluster configuration allows creating only 2 virtual machines, even though the CPU utilization is less when compared to memory utilization. In order to test our monitoring system in a heterogeneous system, VMware Workstation 8 was used to create virtual machine instances and an additional web service application was executed between host and virtual machines along with the same CPU and memory intensive applications were executed and results were discussed in table 2. System with hardware configuration of CPU with 4 x 2.44 GHz, Memory (RAM) with 3.41 GB and local disk with 82 GB, installed windows 7 was considered. The memory usage increases drastically when compared to CPU usage that also limits creating more virtual machines. So, to solve this issue storage consolidation may be considered.

| Number of Virtual Machines and Resource Utilization % | Without virtual Machine instance | With 1 Virtual Machine instance | With 2 Virtual Machine instances | With 3 Virtual Machine instances |
|-------------------------------------------------------|----------------------------------|---------------------------------|----------------------------------|----------------------------------|
|                                                       | CPU Idle % | Memory Usage % | CPU Idle % | Memory Usage % | CPU Idle % | Memory Usage % | CPU Idle % | Memory Usage % |
| Web service application                                | 95         | 21             | 92          | 46              | 90          | 65              | 88          | 82              |
| CPU intensive application                              | 93         | 24             | 85          | 45              | 82          | 68              | 80          | 84              |
| Memory Intensive application                           | 94         | 30             | 87          | 49              | 83          | 71              | 81          | 87              |
From our results it is clear that, a processor with i5-2450M CPU @ 2.50GHz speed, 4GB RAM and 64 bit operating system, allows creating maximum of 3 virtual machine instances. If we consider any high end server, such as Intel Xeon (6 core, 2.93 GHz) with 2 processors, 16GB RAM, minimum 3x146 GB Raid level 5 hard disk, maximum of 12 virtual machines may be created which is an ongoing research issue among cloud researchers. So, in a real data center, there are many such servers running with virtual machines and applications running on these virtual machines need to be migrated either to the same physical server or to different physical server in order to optimize the CPU and memory utilization.

The Linux cluster is analyzed for an hour with virtual machine instances using Ganglia monitoring system for overall system load, memory usage, and CPU and network utilization and projected in graphical form. The physical machine consists of 4 processors, a single node and their corresponding load/processor is shown in Fig. 2. (a). The memory capacity is limited to 3.41 GB which limits creating more than 2 virtual machines is shown in Fig. 2. (b). Fig. 2. (c). shows the network utilization. The graph shows high network traffic at the end. Even though CPU is utilized less than 25%, because of limited RAM capacity the system is not able to create more virtual machines is depicted in Fig. 2. (d). this can be overcome by enforcing scalability in our monitoring system which will be implemented in our future work.

![Fig. 2. (a) Linux_Cluster load/processors](image1)

![Fig. 2. (b) Linux_Cluster Memory usage](image2)

![Fig. 2. (c) Linux_Cluster Network usage](image3)

![Fig. 2. (d) Linux_Cluster CPU usage](image4)
The Linux cluster has a frontend node over which virtual machines are created. Fig. 2 (e) shows the overall utilization of CPU by frontend node that has virtual machines installed. Fig. 2 (f) shows the memory utilized by frontend node. On comparing the utilization of CPU and memory in frontend node, CPU is utilized is less than the memory.

5. Conclusion

Cloud providers must ensure that end user receives the services with reliable and optimal business experience. It is essential to monitor the actual resource usages to avoid over and under estimation of resource levels. So, in this paper, architecture is presented for resource monitoring using agent based system. Since, agents can be loaded anywhere it is easy to use in a real cloud environment for monitoring purposes. As part of our future work, we would like to work on mobile agents that take automatic decisions based on resource usage statistics so that it is beneficiary to both the Cloud Provider and Cloud Consumer.

References

[1] http://csrc.nist.gov/groups/SNS/cloud-computing/cloud-diet-15.doc
[2] John W.Rittinghouse, Janes.F.Ransome, Cloud Computing Implementation, Management and security, CRC Press.
[3] Don Jones, Definitive Guide to Monitoring the Data Center, Virtual Environments, and the Cloud, Nimsoft: from the data center to the Cloud. Real time publishers.
[4] Giuseppe Aceto, Alessio Botta, Walter de Donato, Antonio Pescapè, Cloud Monitoring: A survey, Computer Networks, Elsevier 2013 – Article in press
[5] Shicong Meng, Ling liu, Ting Wang, “State Monitoring in cloud Data Centers”. IEEE transactions on Knowledge and Data Engineering. Vol 23. No.9 September 2011.
[6] Gregory Katsarios, George Kousiouris, Spyridon V. Gogouviits, Dimosthenis Kyriazis, Andreas Menychtas, Theodora Varvarigou, “ A Self-adaptive hierarchical monitoring mechanism for clouds. Journal of Systems and Software 8 December 2011.
[7] Atsuto Inomata, Taiki Morikawa, Minoru Ikebe, Yoshihiro Okamoto, Satoru Noguchi, Kazutoshi Fujikawa, Hideki Sunahara SK, Md, Mizanur Rahman "Proposal and Evaluation of a Dynamic Resource Allocation Method based on the Load of VMs on IaaS, 2011."
[8] Jaideep Moses, Ravi Iyer, ramesh illikal, sadaopan srnivasan, Konstanitonos Aisopos, Shared Resource Monitoring and Throughput Optimization in cloud computing datacenter.

[9] Keijiang Ye, Xiaohong Jiang, Dawei Huang chen, Bei wang, Live Migration of Multiple virtual machines with resource reservation in Cloud computing envirments. IEEE 4th international conference on cloud computing. 2011.

[10] Johan Tordsson, Ruben S.Montero, Rafael Moreno, and Ignacio.M.Llorente, cloud brokering mechanism for optimized placement of virtual machines across multiple providers, Future generation computer systems 28(2012) 358-367.

[11] Mohit Dhingra, J. Lakshmi, S. K. Nandy, Resource Usage Monitoring in Clouds, ACM/IEEE 13th International Conference on Grid Computing. 2012.

[12] Gihun Jung, Kwang Mong Sim, Agent-based Adaptive Resource Allocation on the Cloud Computing Environment, International Conference on Parallel Processing Workshops, 2011.

[13] Qi Liu Georgios K. Theodoropoulos, Dilma Da Silva Elvis S. Liu, Towards An Agent-Based Symbiotic architecture for Autonomic management Of Virtualized Data Centers, Proceedings of the 2012 Winter Simulation Conference, C. Laroque, J. Himmelspach, R. Pasupathy, O. Rose, and A.M. Uhrmacher, eds

[14] Karim Benzidane, Saad Khoudali, Abderrahim Sekkaki, Autonomous Agent-based Inspection for inter-VM Traffic in a Cloud Environment, The 7th International Conference for Internet Technology and Secured Transactions (ICITST-2012)

[15] Ayalew Belay Habit, Cellular-Cloud Integration Framework in Support of Real-Time Monitoring and Management of Traffic on the Road: the Case of Ethiopia, MEDES’12, October 28-31, 2012, Addis Ababa, Ethiopia. 2012 ACM

[16] Kun Ma, Runyuan Sun, Ajith Abraham, Toward a lightweight framework for monitoring public clouds, 978-1-4673-4794-5/12 2012 IEEE