INTERNATIONAL JOURNAL OF RESEARCH IN COMPUTER APPLICATIONS AND ROBOTICS
www.ijrcar.in
Vol.2 Issue.1, Pg.: 139-144
January 2014

INTERNATIONAL JOURNAL OF RESEARCH IN COMPUTER APPLICATIONS AND ROBOTICS
ISSN 2320-7345

DYNAMIC RESOURCE ALLOCATION FOR GREEN CLOUD COMPUTING

P. Malathi[1], R. Kanaga Selvi[2],
PG Student[1], Assistant Professor[2],
Department of Computer Science and Engineering,
Sree Sowdambika College of Engineering,
Aruppukottai, Virudhunagar, TamilNadu, India,
lathi90@gmail.com, rkanagaselvi@gmail.com.

ABSTRACT

Cloud Computing allows users to use resources based on the needs of the corresponding applications. The one of the most important technique in the cloud computing is Virtualization technique which is used for the multiplexing of resources, servers, etc. In this proposal, the system that uses Virtualization technique to allocate resources dynamically based on the demands. It supports green computing. The overall utilization of server resources is controlled by Resource Management System of SPAR method. Experiments show that the performance of our method is better than the existing methods in most situations.

Index terms - Allocating Resources, Balancing Load, Cloud Computing, Dynamic Utilization, Green Computing, Instances, Virtualization.

INTRODUCTION

Cloud computing allows transformational changes with usability, performance, elasticity and security over measurement and characterization for load prediction. But based on the specifications from the cloud users it cannot be modeled accurately based on raw measures which show little variability and less auto-correlation . Virtualization is the one of the most important techniques used in Cloud computing. It supports the cost efficiency while the usage of resources, on demand services and provides resource scalability. The comparison between traditional data center oriented models over the enhanced models and computational services is satisfying the users by providing the quality of service (QoS). Towards to provide the optimum in both computing environment and resources used must give the efficiency, reliability with the co-ordination of the above. This requires the delivery of a set of virtual resources, dynamically allocated to the corresponding server within networked clouds. Pay-per-use infrastructure method is the main advantage of cloud computing. Because it is used to host hundreds of thousands of applications to face the challenges in the resource utilization, resource management, resource pre-reservation [1], [2]. So with the vast positive advantages of Virtualization is becoming most popular technology to frame the infrastructures in the virtualized cloud environment. Among all the Virtual Machine Monitors (VMMs), Xen is defined to be a conspicuous hypervisor based VMM [3]. This VMM is actually used to provide the communication between the Virtual machine (VMs) and Physical Machines (PMs). This communication is hidden from the cloud users. In such cases the each PM must have the sufficient resources which will be requested in the future by the VMs. The mapping between the above two machines is possible means, then the mapping is migrated using the migration list in the VM in order to support green computing by minimizing the number of PMs [4], [5]. Cloud computing allows Transformational changes with usability, performance, elasticity and security over measurement and characterization for load prediction. But based on the specifications from the cloud users it cannot be modeled accurately based on raw measures which show little variability and less auto-correlation [6].

In this paper, we presented the design and implementation of dynamic resource allocation in the Virtualized Cloud Environment which maintains the balance between the following two goals.

Goals to achieve:

Overload Avoidance. The capacity of a PM must satisfy the resource needs from all VMs running on it. Or else, the PM is overloaded and leads to provide less performance of its VMs.

Green computing. The number of PMs used should be optimized as long as they could satisfy the needs of all VMs. And Idle PMs can be turned off to save energy.

There is an in depth trade off between the two goals in the face of changing resource needs from all VMs. To avoid the overload, should keep the utilization of PMs low to reduce the possibility of overload in case the resource needs of VMs increase later. For green computing, should keep the utilization of PMs reasonably high to make efficiency in energy [7].

SYSTEM OVERVIEW

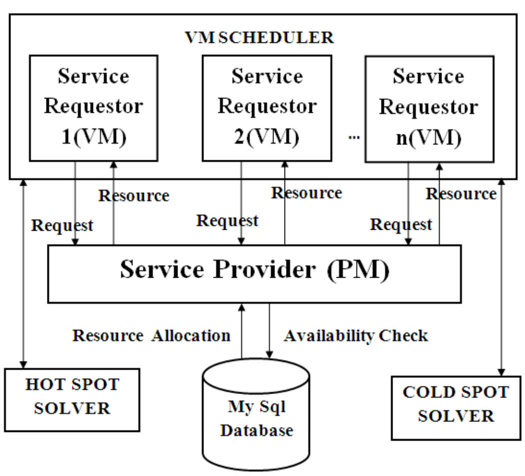

Figure 1. System Architecture

The architecture of the system is represented in Figure 1. The VM scheduler contains a number of VMs. Each Service requestor (VM) encapsulates one or more applications such as Webserver, DNS, Map/Reduce, etc. The specifications of each VMs are sent to the corresponding PM. Then the PM checks for the availability of the resource and checks the memory availability. If there is a possibility to allocate the respective resource with no violation in memory means the dynamic resource allocation is achieved in the VM. The MySql database contains all the data exchanged between all the pair of VM and PM. It updates itself when any changes occur in the communication between PM and VM. The Service provider (PM) communicates with the database and number of clients in the Virtual Scheduler based on the needs of the application.

ALGORITHM

All the conventional algorithms from being utilized for the placement and routing the systems in the cloud network. An implementation in which the algorithm is executed in parallel on separate computers and intra computer. The method we introduced here is SPAR. The partitions of SPAR method are hot spot, cold spot. The hotspot solver connected with the VM scheduler detects if any PM's rate of resource utilization is above the hot threshold (i.e., *hot spot*). If, so the VMs running on that corresponding PM will be reduce its load by implementing the SPAR technique. By implementing the above technique the system achieves "Load balancing". Similarly the cold spot solver checks if the rate of resource utilization is below the green computing threshold (i.e., *cold spot*). If so, those PMs will be turned off in order to save energy or those service providers will be utilized by the any one of the service requesters. By implementing the above methodology the system achieves "Green computing". To avoid the unevenness of resource utilization in different service providers, here we introduce the concept of "Skewness" in order to measure the heterogeneous resources. By minimizing skewness we can improve the overall resource utilization. This concept is associated with the following formula:

Skewness(i) = $\sqrt{\sum_{j=1}^{n}(\frac{r_j}{r}-1)^2}$, Where r_j is the average utilization of all resources for server i. In practice, not all the heterogeneous resources are performed critically and hence we need to consider the bottleneck resources in the above calculations. This is a method for improving memory efficiency. It helps to remember something for long term usage.

The technique executed periodically to evaluate the resource allocation table status. We defined a secondary server of *hotspot* if the utilization of any of its resources is above a *hot threshold* . This indicates that the there is a corresponding server is overloaded and the VM running on it should avoid the overload using the formula for the hot threshold:

Hot threshold (h) = $\sum_{r\in R}(r-r_t)^2$, Where R is the set of overloaded resources in a server I and r_j are the hot spot for the resource r. We defined the hotspot (h) as the square number of its resource utilization beyond the above formula. Actually these overloaded resources are considered only for the calculation.

We defined the cold spot when the utilization of a host is below the cold threshold. This indicates that the particular host will be turned off to save energy or used for any other operation. This cold spot is also named as a green computing threshold. However, we manage that the host must use actively even it has one VM running on it. Finally, we defined a warm threshold to be a balanced level of resource utilization that is sufficient to provide the highest performance.

EXPERIMENTAL RESULT

Our experiments are conducted using a group of 10 Lenovo hosts with Intel dual core and 2GB of RAM. The system runs on Windows XP operating system. We update the resource allocation by the Allocation table. The hosts are connected in LAN (Local Area Network). The server host contains the storage of the updated dataset. All the client hosts can access that dataset through server host.

i) Memory Efficiency:

We started our evaluation to achieve the effectiveness of the method used here in overload avoidance and green computing. We start with a small scale experiment consisting of one PM and two VMs. By this connection setup we can present the results for all servers in Figure 3.

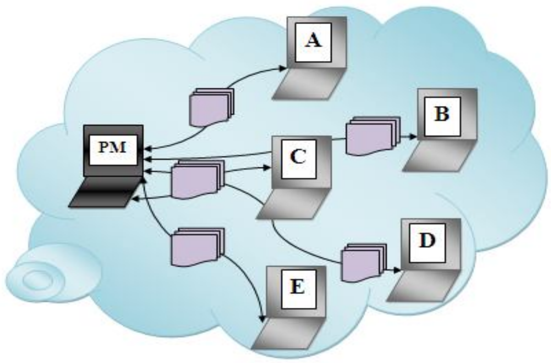

Figure 2. Cloud environment Establishment

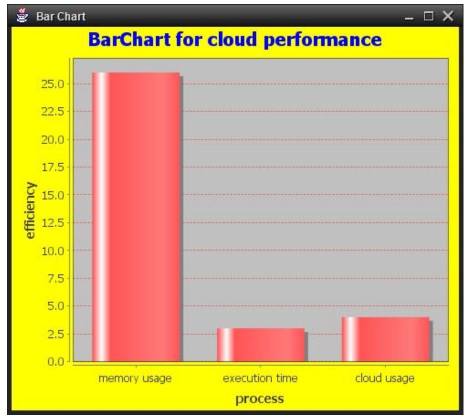

Figure 3. Efficiency Graph.

Here the memory utilization is high. But when we combine the number VMs the memory utilization is distributed among a set of VMs. The increasing number of VMs will be implemented in the future in order to increase the memory efficiency (Figure 2) associated with the increased usage of cloud in a variable execution time on demand.

ii) Resource Balance:

The goals of this paper are balanced by the resource management method we introduced here. i.e., By minimizing the number of hosts in the cloud environment the memory utilization may be high. But in case of maximizing the number of VMs the overload can be avoided by distributing the load in the connected VMs. In the above Figure the efficiency is plotted over memory ,cloud usage, execution time. The efficiency can scale up or scale down. So we aim to balance the goals which were mentioned in the Section I by means of providing the best balance between the goals. This balancing implementation is will be achieved in later. And this achievement supports green computing.

RELATED WORK

i) Online Load Balancing:

Even for the websites of e-commerce sites are not clearly known by the users. To access that website the user must know how the various servers will handle the load in online. Actually, this type of online loads, performance is characterized by the HTTP requests [9]. So we can implement this system for the load balancing concept of online application.

ii) Power Management in VM:

The modern data centers have two issues such as Better Performance and the management of power in the Virtualized environment. The challenging complexity of application with balanced workload and shared Virtualized infrastructure. [10] [11]. Our goal is to develop a technique similar to the GNP which deals with the Virtualized dynamic environments in the efficient manner based on the connections established between the PMs and VMs. The limitations of GNP are as follows: 1) The environment composed a of number of VMs with the PMs which makes the communication easier than other methods. 2) High degree of Reusability [12]. Energy consumption in the cloud data centers is variation in its performance. The effective and smart utilization of virtual machines (VMs) in Cloud computing offers the service providers an opportunity to combine the various workloads from fewer server hosts, finally it improves the resource utilization as well as eliminating the power consumption by the idle machine. Dynamic Frequency Voltage Scaling (DVFS) technology, which is extensively supported by the server manufacturers [15].

iii) Resource Balance:

The process of allocating resources in a dynamic way is an effective way to improve the resource utilization as well as the energy efficiency in cloud data centers. Determining the best reallocation of VMs when overload occurs in order to provide the Quality Of Service (QoS). Based on the implementation of the Markov Chain Model the problem of overload is identified and the optimal solutions are achieved [13]. Dynamically allocating resources is the one of the important issues in a cloud environment. Particularly some of the service requesters have the Service Level Agreement (SLAs). So the overall profit of that environment depends on how much amount SLAs of the systems meet [14].

CONCLUSION

We have presented the design, implementation, and evaluation virtualized cloud environment for dynamic resource allocation. This system multiplex the virtual machines based on the demands of the end users. We use SPAR method to provide memory efficiency in the servers. Finally this system achieves green computing and load balancing between heterogeneous end users.

ACKNOWLEDGMENT

I would like to express my deep gratitude and respect to Mrs. R. Kanaga Selvi whose advices and insight was valuable to me. She inspired me to continue the development of my Post Graduate Paper work. The authors would like to thank the anonymous reviewers for their valuable feedback.

REFERENCES

[1] Papagianni.C., Leivadeas.A., Papavassiliou.S., Maglaris.V., Cervello-Pastor.C., Monje.A., "On the optimal allocation of virtual resources in cloud computing networks", Computers, IEEE Transactions on (Volume:62 , Issue: 6), June. 2013.

[2] Yuan Yuan, Wen-Cai Liu, "Efficient resource management for cloud computing", System Science, Engineering Design and Manufacturing Informatization (ICSEM), 2011 International Conference on (Volume:2), October. 2011.

[3] P. Barham, B. Dragovic, K. Fraser, S. Hand, T. Harris, A. Ho, R.Neugebauer, I. Pratt, and A. Warfield, "Xen and the Art of Virtualization," Proc. ACM Symp. Operating Systems Principles (SOSP '03), Oct. 2003.

[4] C. Clark, K. Fraser, S. Hand, J.G. Hansen, E. Jul, C. Limpach, I.Pratt, and A. Warfield, "Live Migration of Virtual Machines," Proc. Symp. Networked Systems Design and Implementation (NSDI '05), May 2005.

[5] M. Nelson, B.-H. Lim, and G. Hutchins, "Fast Transparent Migration for Virtual Machines," Proc. USENIX Ann. Technical Conf., 2005.

[6] Saripalli, P., Kiran, G.V.R., Shankar, R.R., Narware, H., Bindal, N. "Load Prediction and Hot Spot Detection Models for Autonomic Cloud Computing" , Utility and Cloud Computing (UCC), Dec. 2011.

[7] Paul Barham, Boris Dragovic, Keir Fraser, Steven Hand, Tim Harris, Alex Ho, Rolf Neugebauer, Ian Pratt, Andrew Warfield," Xen and the Art of Virtualization", University of Cambridge Computer Laboratory, UK, CB3 0FD.

[8] Zhen Xiao, Weijia Song, Qi Chen, "Dynamic Resource Allocation Using Virtual Machines for Cloud Computing Environment", Parallel and Distributed Systems, IEEE Transactions on (Volume:24 , Issue:6), June. 2013.

[9] Menasce, D. "TPC-W: a benchmark for e-commerce", Internet Computing, IEEE Volume: 6 , Issue: 3 , 2002.

[10] YanfeiGuo, Xiaobo Zhou, "Coordinated VM Resizing and Server Tuning: Throughput, Power Efficiency and Scalability", Modeling, Analysis & Simulation of Computer and Telecommunication Systems (MASC OTS), IEEE 20th International Symposium on 2012.

[11] Pedram, M., Inkwon Hwang," Power and Performance Modeling in a Virtualized Server System" , Parallel Processing Workshops (ICPPW) 39th International Conference on 2010.

[12] Shingo Mabu, Kotaro Hirasawa, Jinglu Hu, **"**A Graph-Based Evolutionary Algorithm: Genetic Network Programming (GNP) and Its Extension Using Rinforcement Learning**",** Massachusetts Institute of Technology, August. 2007.

[13] Beloglazov.A., Buyya.R., "Managing Overloaded Hosts for Dynamic Consolidation of Virtual Machines in Cloud Data Centers under Quality of Service Constraints", Parallel and Distributed Systems, IEEE Transactions on Volume: 24 , Issue: 7 ,2013.

[14] Yanzhi Wang, Shuang Chen, Pedram, M. "Service Level Agreement-Based Joint Application Environment Assignment and Resource Allocation in Cloud Computing Systems", Green Technologies Conference IEEE, 2013.

[15] Hailong Yang, Qi Zhao, Zhongzhi Luan, Depei Qian, "UVMPM: A Unitary Approach for VM Power Metering Based on Performance Profiling", Parallel Processing Workshops (ICPPW), 2012.