A Survey on Auto Live Migration Mechanism in Cloud Environment

Mayuri B. Katosana
Department of Computer Engineering
L. D. College of Engineering
Ahmedabad, India
mkatosana.mk@gmail.com

Prof. Hetal A. Joshiyara
Department of Computer Engineering
L. D. College of Engineering
Ahmedabad, India
hetaljoshiyara@gmail.com

***Abstract*—**Cloud Computing has recently emerged as a compelling paradigm for delivering computing services to users as utilities in a pay-as-you-go manner over the internet. Virtualization is a key concept in cloud computing. Virtualization technology refers to the creation of a virtual machine that acts like a real hardware with an operating system. Live migration is the task of moving a virtual machine from one physical hardware environment to another without disconnecting the client. Its facilities for efficient utilization of resources (CPU, memory, Storage) to manage load imbalance problem and also useful for reduction in energy consumption and fault management. For live migration of the virtual machine cloud provider needs to monitor the resources of all hosts continuously. So there are techniques for automation of this live migration when required. This method is called auto live migration techniques. This paper presents a detailed survey on Auto Live Migration of Virtual machines (VM) in cloud computing.

Keywords*-Cloud Computing; Load Balancing; Virtual Machine; Live Migration; Resource Utilization; Auto Live Migration*

*****

I. INTRODUCTION

Cloud computing is a method of providing services to the consumer by using internet through web-based technology and applications. Cloud computing offers services like infrastructure (IaaS-Infrastructure as a Service), platform (PaaS-Platform as a service) and application or software (SaaS-Software as a Service). It is made available services to the user in a Pay-As-You-Go manner. Some examples are Microsoft Azure, Google App Engine and Amazon Web Service, etc.

Cloud computing provides scalable, on-demand, cost-effective, device independent and reliable services to its user. Traditionally, enterprises purchased their computing infrastructures to run a real-time application, such as financial analysis, distributed data processing, real-time databases, etc. They can access their files and data on any computer or at any place through the Internet. A standard definition for cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, server, storage, application, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.

Virtualization is a core concept in data centers. Virtualization refers to the creation of virtual machine on a physical machine. A single physical machine can have many smaller virtual machines in which isolated operating system instances are running. The machine on which virtual machine is created is called a host machine, and the virtual machine is referred as a guest machine. This created virtual machine managed by software known as a hypervisor. The hypervisor is a low-level program or software which acts as a Virtual Machine Manager. KVM, Xen, Microsoft Hyper-V and VMware ESX are popular virtualization software.

In cloud computing because of the fickle nature of the cloud users, all cloud hosts do not have an equal amount of load. So it will degrade the performance of the system. Load balancing is a key technology to maintain the imbalance load of a particular host. It helps to distribute workload across multiple hosts to ensure that no single host gets overloaded. It is also helpful for proper utilization of resources. There are various types of load possible in cloud computing like CPU, memory, storage and network load. Load balancing is the technique of finding an overloaded host and transferring the extra load to another host.

The question is that how to achieve effective load balance, how to improve resource utilization and how to schedule resources in cloud computing environment. The solution to this problem is live migration technique to balance the load between hosts. Live migration means transferring the virtual machine from one host to another without disconnecting the clients and applications. Live migration is a key technique to maintain the resource dynamically.

Now it is difficult to monitor all host's and virtual machine's resource statistics. So here comes the concept of auto live migration. In auto live migration cloud itself or with any third party application monitors all resources of the cloud hosts and does live migration when needed. There are so many algorithms for deciding the right source host, destination host and time to migrate.

In this paper, we survey on the performance technologies of the VM live migration techniques. We survey the literature on the evaluation of various VM auto live migration techniques

and identify the performance metrics. All the existing auto live virtual machine migration techniques are studied and classified based on these metrics. This paper is organized as follows. Section II gives a brief introduction of Virtual Machine Migration (VMM). Auto live VM Migration techniques are surveyed in section III. Live Migration related performance metrics described in section IV. Section V gives benefits of auto live migration. We conclude our work in section VI.

II. OVERVIEW OF VIRTUAL MACHINE MIGRATION

Virtualization technology provides facility migrate VM (Virtual machine) from one host to another host. Virtualization is a technique, which allows multiple operating systems run concurrently on the single physical machine. Virtual machines are handled by hypervisor which is also known as virtual machine manager. Live migration of virtual machine is helpful for load balancing, energy saving, efficient resource utilization.

Virtual Machine Migration methods are divided into two types: Hot (live) migration and cold (non-live) migration. The status of the VM loses, and the user can notice the service interruption in cold migration. The virtual machine keeps running while migrating and does not lose its status, so the user doesn't feel any interruption in service in hot (live) migration. In live migration process, the state of a virtual machine to migrate is transferred. The state consists of its memory contents and local file system.

III. AUTO VM MIGRATION TECHNIQUES IN CLOUD

Author approach two algorithms for a different task in their created environment. In first, Migration algorithm which is migrating the overloaded process from one machine to another machine in the same cloud environment. In second, scheduling algorithm which schedule the processes in consideration of the resource utilization that is based on memory usage, CPU usage, and throughput and processing time. The main aim of this paper is to allocate requests resources by using process migration algorithm and schedules effectively the arrived process by using effective scheduling algorithm to get maximum resource utilization [1].

In this paper author discussed live migration technique that provides the ability of dynamic virtual machine consolidation using adaptive utilization threshold based on CPU usage prediction which can easily manage the high level of SLA and reduces the number of VM migration in between the host. This approach describes the phenomenon of the underload and overload detection whenever the host has underload all the virtual machines from the host migrated to another host, and the particular host is switched off for the time being. Moreover, the requested amount CPU remand exceeds the available capacity on that host then this host considered as an overloaded host and some VMs migrate to another host for maintaining the SLA level [2].

The author proposed Autonomous Network Aware VM migration Strategy that observes the current demand level of network and performs appropriate action based on what it is experiencing. The author also proposed a dynamic Reinforcement Learning (RL) network aware approach, enabling a single RL agent to learn the most opportune time to migrate a group of VM's, Depending upon the current utilization of cloud's network resource. This autonomous approach is sensitive to high demand request to a network resource and makes efficient use of resource available [3].

The author presented dynamic resource allocation and energy saving algorithms. To implement these two algorithms with live migration of virtual machines, the author builds an infrastructure platform based on cloud software OpenStack. They use the Power Distribution Unit (PDU) to monitor system status and record power consumption. The result of the PDU power consumption records verifies that the two proposed algorithms are effective [4].

In this paper author proposed Dynamic Weighted Live Migration (DWLM) to manage load Imbalance problem. DWLM algorithm checks the VMs availability and available jobs in a queue which are not allocated in VMs due to overloaded condition on VMs. DWLM manages the load and finds the VMs and migrate the jobs with the help of threshold model. Migration and Utilization part depends on the threshold value. Threshold value helps to separate the high and low load VMs. The proposed mechanism results outcomes compared with Migration time, Scalability, Throughput and availability factor [5].

The author presented pull strategy and push strategy to balance the load in the system with multiple VM through automated live migration. When the push strategy is used, overloaded host tries to migrate workload to less loaded hosts. On the other hand, the pull strategy can quickly re-distribute the load of the system when the load is in the range of low-to-medium [6].

Author-defined algorithm to solve traditional migration problem that needs manually determining whether the machine load is too heavy or not. In this scheme, virtual machine load level is automatically detected. The author proposed Automatic Dynamic Migration to balance resources in the OpenStack environment. The author also planned in this paper to integrate algorithm with resource monitoring open source software for real-time monitoring of resources. Thus when load imbalance of resources occurs, the system automatically performs live migration to manage the load on the system [7].

In this article, the author proposed Multiple Regression Host Detection (MRHOD) algorithm that uses CPU utilization, memory utilization and bandwidth utilization for host overload detection using hybrid factors to enhance VM consolidation. This algorithm reduces energy consumption while ensuring a high level of adherence to SLA. In this algorithm utilization of resources based on three parameters (CPU, Memory, Bandwidth) instead of one parameter only (CPU). This approach reduces power consumption by six times compared to single factor algorithm using random workload. They also developed Hybrid Local Regression Host Overload Detection algorithm (HLRHOD) that is based on local regression using hybrid factors [8].

The author presented a CPU usage prediction method based on the linear regression technique. The linear regression is a strong statistical method that used in machine learning schemes to estimate a prediction function. The linear regression creates the function according to the past utilization values in a host. The function can forecast the future short- term utilization based on the current requested utilization in each host. If the prediction utilization is greater than the current utilization capacity, then the host will be overloaded. Therefore, migration of virtual machine starts [9].

In this paper, a Dynamic cloud environment Virtual Machine Migration algorithm VM-DFM (Virtual Machine Dynamic Forecast Migration) proposed. This algorithm reduces the amount of virtual machine migration. VM-DFM model of cloud environment periodic virtual memory consumption data statistics and forecast, the algorithm can be applied to excessive memory consumption load in cloud environment of physical nodes to the appropriate VM live migration [10].

The author presented OpenStack Neat to provide an extensible framework for dynamic consolidation of VMs based on the OpenStack platform. The functionality covered by this project will be implemented in the form of services separate from the core OpenStack services. The services of this project will interact with the core OpenStack services using their public APIs. The author also proposed a benchmark suite for evaluating and comparing dynamic VM consolidation algorithms. The proposed benchmark suite comprises OpenStack Neat as the base software framework, a set of real-world workload traces, performance metrics and evaluation methodology [11].

IV. PERFORMANCE METRICS

The following metrics are usually used to measure the performance of live migration:

*a) Downtime:*The time during which the migrating VM's is not executing. It includes the transfer of processor state.

b) Total Energy Consumption: This is defined as the sum of energy consumed by the physical resources of a data center as a result of application workloads.

c) Total Migration Time: This is the total time from start of migration to finish. Total Migration Time is essential because it affects the release of resources on both participating nodes as well as within the VMs.

d) Application Degradation:: This is the extent to which migration slows down the applications executing within the VM.

e) SLA (Service Level Agreement): It can be determined regarding such characteristics as minimum throughput or maximum response time delivered by the deployed system.

V. BENEFITS OF AUTO LIVE MIGRATON

Auto Live Migration provides load balancing of overloaded or under loaded hosts. So it efficiently utilizes the resources of the system. Because of the better utilization, Auto Live Migration reduces the power consumption. And also the Auto Live Migration makes system fault tolerant by managing the load of the system.

VI. CONCLUSION

This paper is a survey of auto live migration of virtual machine techniques. Auto live migration does load balancing of cloud system without the need for user interference. The key concept of auto live migration is to choose an overloaded or going to overload host and do migration such that SLA of any instance does violate. In this paper, we have described the latest auto live migration mechanism.

REFERENCES

[1] A. Upadhyay and P. Lakkadwala, "Migration of over loaded process and schedule for resource utilization in Cloud Computing," 2015 4th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO) (Trends and Future Directions), Noida, India, 2015, pp. 1-4.

[2] Vikas Malik and C R Barde. Article: Live Migration of Virtual Machines in Cloud Environment using Prediction of CPU Usage. International Journal of Computer Applications 117(23):1-5, May 2015.

[3] Duggan, Martin & Duggan, Jim & Howley, Enda & Barrett, Enda. (2016). An Autonomous Network Aware VM Migration Strategy in Cloud Data Centres. . 10.1109/ICCAC.2016.9.

[4] C. C. Chen, P. L. Sun, C. T. Yang, J. C. Liu, S. T. Chen and Z. Y. Wan, "Implementation of a Cloud Energy Saving System with Virtual Machine Dynamic Resource Allocation Method Based on OpenStack," 2015 Seventh International Symposium on Parallel Architectures, Algorithms and Programming (PAAP), Nanjing, 2015, pp. 190-196.

[5] P. K. Tiwari and S. Joshi, "Dynamic weighted virtual machine live migration mechanism to manages load balancing in cloud computing," 2016 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), Chennai, 2016, pp. 1-5.

[6] Mattias Forsman, Andreas Glad, Lars Lundberg, Dragos Ilie, Algorithms for automated live migration of virtual machines, In Journal of Systems and Software, Volume 101, 2015, Pages 110-126, ISSN 0164-1212.

[7] C. T. Yang, Y. T. Liu, J. C. Liu, C. L. Chuang and F. C. Jiang, "Implementation of a Cloud IaaS with Dynamic Resource Allocation Method Using OpenStack," 2013 International Conference on Parallel and Distributed Computing, Applications and Technologies, Taipei, 2013, pp. 71-78.

[8] Abdelsamea, A., A. A. El-Moursy, E. E. Hemayed, and H. Eldeeb, "Virtual machine consolidation enhancement using hybrid regression algorithms", Egyptian Informatics Journal: Elsevier, 2017.

[9] F. Farahnakian, P. Liljeberg and J. Plosila, "LiRCUP: Linear Regression Based CPU Usage Prediction Algorithm for Live Migration of Virtual Machines in Data Centers," 2013 39th Euromicro Conference on Software Engineering and Advanced Applications, Santander, 2013, pp. 357-364.

[10] J. Chen, Y. Qin, Y. Ye and Z. Tang, "A Live Migration Algorithm for Virtual Machine in a Cloud Computing Environment," 2015 IEEE 12th Intl Conf on Ubiquitous Intelligence and Computing and 2015 IEEE 12th Intl Conf on Autonomic and Trusted Computing and 2015 IEEE 15th Intl Conf on Scalable Computing and Communications and Its Associated Workshops (UIC-ATC-ScalCom), Beijing, 2015, pp. 1319- 1326.

[11] Beloglazov, Anton & Buyya, Rajkumar. (2014). OpenStack Neat: A framework for dynamic and energy-efficient consolidation of virtual machines in OpenStack clouds. Concurrency and Computation: Practice and Experience. 27. . 10.1002/cpe.3314.